

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:22:42 ; Search time 41.23 Seconds
(without alignments)
4212.634 Million cell updates/sec

Title: US-09-767-215-2
Perfect score: 5149
Sequence: 1 MGELCRRDSALTADEETLW.....VROIAADEQKKVVWTEOSPR 1004

Scoring table: BL0SUMM62
Gappen 10.0 , capext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 5622222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : SPTREMBL19:
 1: sp_archaea:
 2: sp_bacteria:
 3: sp_fungi:
 4: sp_human:
 5: sp_invertebrate:
 6: sp_mammal:
 7: sp_mhc:
 8: sp_organelle:
 9: sp_phage:
 10: sp_plant:
 11: sp_rabbit:
 12: sp_virus:
 13: sp_vertebrate:
 14: sp_unclassified:
 15: sp_virus:
 16: sp_bacteriophage:
 17: sp_archaea:
 18: sp_bacteria:
 19: sp_fungi:
 20: sp_human:
 21: sp_invertebrate:
 22: sp_mammal:
 23: sp_mhc:
 24: sp_organelle:
 25: sp_phage:
 26: sp_plant:
 27: sp_rabbit:
 28: sp_virus:
 29: sp_vertebrate:
 30: sp_unclassified:
 31: sp_virus:
 32: sp_bacteriophage:
 33: sp_bacteria:
 34: sp_fungi:
 35: sp_human:
 36: sp_invertebrate:
 37: sp_mammal:
 38: sp_mhc:
 39: sp_organelle:
 40: sp_phage:
 41: sp_plant:
 42: sp_rabbit:
 43: sp_virus:
 44: sp_vertebrate:
 45: sp_unclassified:
 46: sp_bacteriophage:
 47: sp_bacteria:
 48: sp_fungi:
 49: sp_human:
 50: sp_invertebrate:
 51: sp_mammal:
 52: sp_mhc:
 53: sp_organelle:
 54: sp_phage:
 55: sp_plant:
 56: sp_rabbit:
 57: sp_virus:
 58: sp_vertebrate:
 59: sp_unclassified:
 60: sp_bacteriophage:
 61: sp_bacteria:
 62: sp_fungi:
 63: sp_human:
 64: sp_invertebrate:
 65: sp_mammal:
 66: sp_mhc:
 67: sp_organelle:
 68: sp_phage:
 69: sp_plant:
 70: sp_rabbit:
 71: sp_virus:
 72: sp_vertebrate:
 73: sp_unclassified:
 74: sp_bacteriophage:
 75: sp_bacteria:
 76: sp_fungi:
 77: sp_human:
 78: sp_invertebrate:
 79: sp_mammal:
 80: sp_mhc:
 81: sp_organelle:
 82: sp_phage:
 83: sp_plant:
 84: sp_rabbit:
 85: sp_virus:
 86: sp_vertebrate:
 87: sp_unclassified:
 88: sp_bacteriophage:
 89: sp_bacteria:
 90: sp_fungi:
 91: sp_human:
 92: sp_invertebrate:
 93: sp_mammal:
 94: sp_mhc:
 95: sp_organelle:
 96: sp_phage:
 97: sp_plant:
 98: sp_rabbit:
 99: sp_virus:
 100: sp_vertebrate:
 101: sp_unclassified:
 102: sp_bacteriophage:
 103: sp_bacteria:
 104: sp_fungi:
 105: sp_human:
 106: sp_invertebrate:
 107: sp_mammal:
 108: sp_mhc:
 109: sp_organelle:
 110: sp_phage:
 111: sp_plant:
 112: sp_rabbit:
 113: sp_virus:
 114: sp_vertebrate:
 115: sp_unclassified:
 116: sp_bacteriophage:
 117: sp_bacteria:
 118: sp_fungi:
 119: sp_human:
 120: sp_invertebrate:
 121: sp_mammal:
 122: sp_mhc:
 123: sp_organelle:
 124: sp_phage:
 125: sp_plant:
 126: sp_rabbit:
 127: sp_virus:
 128: sp_vertebrate:
 129: sp_unclassified:
 130: sp_bacteriophage:
 131: sp_bacteria:
 132: sp_fungi:
 133: sp_human:
 134: sp_invertebrate:
 135: sp_mammal:
 136: sp_mhc:
 137: sp_organelle:
 138: sp_phage:
 139: sp_plant:
 140: sp_rabbit:
 141: sp_virus:
 142: sp_vertebrate:
 143: sp_unclassified:
 144: sp_bacteriophage:
 145: sp_bacteria:
 146: sp_fungi:
 147: sp_human:
 148: sp_invertebrate:
 149: sp_mammal:
 150: sp_mhc:
 151: sp_organelle:
 152: sp_phage:
 153: sp_plant:
 154: sp_rabbit:
 155: sp_virus:
 156: sp_vertebrate:
 157: sp_unclassified:
 158: sp_bacteriophage:
 159: sp_bacteria:
 160: sp_fungi:
 161: sp_human:
 162: sp_invertebrate:
 163: sp_mammal:
 164: sp_mhc:
 165: sp_organelle:
 166: sp_phage:
 167: sp_plant:
 168: sp_rabbit:
 169: sp_virus:
 170: sp_vertebrate:
 171: sp_unclassified:
 172: sp_bacteriophage:
 173: sp_bacteria:
 174: sp_fungi:
 175: sp_human:
 176: sp_invertebrate:
 177: sp_mammal:
 178: sp_mhc:
 179: sp_organelle:
 180: sp_phage:
 181: sp_plant:
 182: sp_rabbit:
 183: sp_virus:
 184: sp_vertebrate:
 185: sp_unclassified:
 186: sp_bacteriophage:
 187: sp_bacteria:
 188: sp_fungi:
 189: sp_human:
 190: sp_invertebrate:
 191: sp_mammal:
 192: sp_mhc:
 193: sp_organelle:
 194: sp_phage:
 195: sp_plant:
 196: sp_rabbit:
 197: sp_virus:
 198: sp_vertebrate:
 199: sp_unclassified:
 200: sp_bacteriophage:
 201: sp_bacteria:
 202: sp_fungi:
 203: sp_human:
 204: sp_invertebrate:
 205: sp_mammal:
 206: sp_mhc:
 207: sp_organelle:
 208: sp_phage:
 209: sp_plant:
 210: sp_rabbit:
 211: sp_virus:
 212: sp_vertebrate:
 213: sp_unclassified:
 214: sp_bacteriophage:
 215: sp_bacteria:
 216: sp_fungi:
 217: sp_human:
 218: sp_invertebrate:
 219: sp_mammal:
 220: sp_mhc:
 221: sp_organelle:
 222: sp_phage:
 223: sp_plant:
 224: sp_rabbit:
 225: sp_virus:
 226: sp_vertebrate:
 227: sp_unclassified:
 228: sp_bacteriophage:
 229: sp_bacteria:
 230: sp_fungi:
 231: sp_human:
 232: sp_invertebrate:
 233: sp_mammal:
 234: sp_mhc:
 235: sp_organelle:
 236: sp_phage:
 237: sp_plant:
 238: sp_rabbit:
 239: sp_virus:
 240: sp_vertebrate:
 241: sp_unclassified:
 242: sp_bacteriophage:
 243: sp_bacteria:
 244: sp_fungi:
 245: sp_human:
 246: sp_invertebrate:
 247: sp_mammal:
 248: sp_mhc:
 249: sp_organelle:
 250: sp_phage:
 251: sp_plant:
 252: sp_rabbit:
 253: sp_virus:
 254: sp_vertebrate:
 255: sp_unclassified:
 256: sp_bacteriophage:
 257: sp_bacteria:
 258: sp_fungi:
 259: sp_human:
 260: sp_invertebrate:
 261: sp_mammal:
 262: sp_mhc:
 263: sp_organelle:
 264: sp_phage:
 265: sp_plant:
 266: sp_rabbit:
 267: sp_virus:
 268: sp_vertebrate:
 269: sp_unclassified:
 270: sp_bacteriophage:
 271: sp_bacteria:
 272: sp_fungi:
 273: sp_human:
 274: sp_invertebrate:
 275: sp_mammal:
 276: sp_mhc:
 277: sp_organelle:
 278: sp_phage:
 279: sp_plant:
 280: sp_rabbit:
 281: sp_virus:
 282: sp_vertebrate:
 283: sp_unclassified:
 284: sp_bacteriophage:
 285: sp_bacteria:
 286: sp_fungi:
 287: sp_human:
 288: sp_invertebrate:
 289: sp_mammal:
 290: sp_mhc:
 291: sp_organelle:
 292: sp_phage:
 293: sp_plant:
 294: sp_rabbit:
 295: sp_virus:
 296: sp_vertebrate:
 297: sp_unclassified:
 298: sp_bacteriophage:
 299: sp_bacteria:
 300: sp_fungi:
 301: sp_human:
 302: sp_invertebrate:
 303: sp_mammal:
 304: sp_mhc:
 305: sp_organelle:
 306: sp_phage:
 307: sp_plant:
 308: sp_rabbit:
 309: sp_virus:
 310: sp_vertebrate:
 311: sp_unclassified:
 312: sp_bacteriophage:
 313: sp_bacteria:
 314: sp_fungi:
 315: sp_human:
 316: sp_invertebrate:
 317: sp_mammal:
 318: sp_mhc:
 319: sp_organelle:
 320: sp_phage:
 321: sp_plant:
 322: sp_rabbit:
 323: sp_virus:
 324: sp_vertebrate:
 325: sp_unclassified:
 326: sp_bacteriophage:
 327: sp_bacteria:
 328: sp_fungi:
 329: sp_human:
 330: sp_invertebrate:
 331: sp_mammal:
 332: sp_mhc:
 333: sp_organelle:
 334: sp_phage:
 335: sp_plant:
 336: sp_rabbit:
 337: sp_virus:
 338: sp_vertebrate:
 339: sp_unclassified:
 340: sp_bacteriophage:
 341: sp_bacteria:
 342: sp_fungi:
 343: sp_human:
 344: sp_invertebrate:
 345: sp_mammal:
 346: sp_mhc:
 347: sp_organelle:
 348: sp_phage:
 349: sp_plant:
 350: sp_rabbit:
 351: sp_virus:
 352: sp_vertebrate:
 353: sp_unclassified:
 354: sp_bacteriophage:
 355: sp_bacteria:
 356: sp_fungi:
 357: sp_human:
 358: sp_invertebrate:
 359: sp_mammal:
 360: sp_mhc:
 361: sp_organelle:
 362: sp_phage:
 363: sp_plant:
 364: sp_rabbit:
 365: sp_virus:
 366: sp_vertebrate:
 367: sp_unclassified:
 368: sp_bacteriophage:
 369: sp_bacteria:
 370: sp_fungi:
 371: sp_human:
 372: sp_invertebrate:
 373: sp_mammal:
 374: sp_mhc:
 375: sp_organelle:
 376: sp_phage:
 377: sp_plant:
 378: sp_rabbit:
 379: sp_virus:
 380: sp_vertebrate:
 381: sp_unclassified:
 382: sp_bacteriophage:
 383: sp_bacteria:
 384: sp_fungi:
 385: sp_human:
 386: sp_invertebrate:
 387: sp_mammal:
 388: sp_mhc:
 389: sp_organelle:
 390: sp_phage:
 391: sp_plant:
 392: sp_rabbit:
 393: sp_virus:
 394: sp_vertebrate:
 395: sp_unclassified:
 396: sp_bacteriophage:
 397: sp_bacteria:
 398: sp_fungi:
 399: sp_human:
 400: sp_invertebrate:
 401: sp_mammal:
 402: sp_mhc:
 403: sp_organelle:
 404: sp_phage:
 405: sp_plant:
 406: sp_rabbit:
 407: sp_virus:
 408: sp_vertebrate:
 409: sp_unclassified:
 410: sp_bacteriophage:
 411: sp_bacteria:
 412: sp_fungi:
 413: sp_human:
 414: sp_invertebrate:
 415: sp_mammal:
 416: sp_mhc:
 417: sp_organelle:
 418: sp_phage:
 419: sp_plant:
 420: sp_rabbit:
 421: sp_virus:
 422: sp_vertebrate:
 423: sp_unclassified:
 424: sp_bacteriophage:
 425: sp_bacteria:
 426: sp_fungi:
 427: sp_human:
 428: sp_invertebrate:
 429: sp_mammal:
 430: sp_mhc:
 431: sp_organelle:
 432: sp_phage:
 433: sp_plant:
 434: sp_rabbit:
 435: sp_virus:
 436: sp_vertebrate:
 437: sp_unclassified:
 438: sp_bacteriophage:
 439: sp_bacteria:
 440: sp_fungi:
 441: sp_human:
 442: sp_invertebrate:
 443: sp_mammal:
 444: sp_mhc:
 445: sp_organelle:
 446: sp_phage:
 447: sp_plant:
 448: sp_rabbit:
 449: sp_virus:
 450: sp_vertebrate:
 451: sp_unclassified:
 452: sp_bacteriophage:
 453: sp_bacteria:
 454: sp_fungi:
 455: sp_human:
 456: sp_invertebrate:
 457: sp_mammal:
 458: sp_mhc:
 459: sp_organelle:
 460: sp_phage:
 461: sp_plant:
 462: sp_rabbit:
 463: sp_virus:
 464: sp_vertebrate:
 465: sp_unclassified:
 466: sp_bacteriophage:
 467: sp_bacteria:
 468: sp_fungi:
 469: sp_human:
 470: sp_invertebrate:
 471: sp_mammal:
 472: sp_mhc:
 473: sp_organelle:
 474: sp_phage:
 475: sp_plant:
 476: sp_rabbit:
 477: sp_virus:
 478: sp_vertebrate:
 479: sp_unclassified:
 480: sp_bacteriophage:
 481: sp_bacteria:
 482: sp_fungi:
 483: sp_human:
 484: sp_invertebrate:
 485: sp_mammal:
 486: sp_mhc:
 487: sp_organelle:
 488: sp_phage:
 489: sp_plant:
 490: sp_rabbit:
 491: sp_virus:
 492: sp_vertebrate:
 493: sp_unclassified:
 494: sp_bacteriophage:
 495: sp_bacteria:
 496: sp_fungi:
 497: sp_human:
 498: sp_invertebrate:
 499: sp_mammal:
 500: sp_mhc:
 501: sp_organelle:
 502: sp_phage:
 503: sp_plant:
 504: sp_rabbit:
 505: sp_virus:
 506: sp_vertebrate:
 507: sp_unclassified:
 508: sp_bacteriophage:
 509: sp_bacteria:
 510: sp_fungi:
 511: sp_human:
 512: sp_invertebrate:
 513: sp_mammal:
 514: sp_mhc:
 515: sp_organelle:
 516: sp_phage:
 517: sp_plant:
 518: sp_rabbit:
 519: sp_virus:
 520: sp_vertebrate:
 521: sp_unclassified:
 522: sp_bacteriophage:
 523: sp_bacteria:
 524: sp_fungi:
 525: sp_human:
 526: sp_invertebrate:
 527: sp_mammal:
 528: sp_mhc:
 529: sp_organelle:
 530: sp_phage:
 531: sp_plant:
 532: sp_rabbit:
 533: sp_virus:
 534: sp_vertebrate:
 535: sp_unclassified:
 536: sp_bacteriophage:
 537: sp_bacteria:
 538: sp_fungi:
 539: sp_human:
 540: sp_invertebrate:
 541: sp_mammal:
 542: sp_mhc:
 543: sp_organelle:
 544: sp_phage:
 545: sp_plant:
 546: sp_rabbit:
 547: sp_virus:
 548: sp_vertebrate:
 549: sp_unclassified:
 550: sp_bacteriophage:
 551: sp_bacteria:
 552: sp_fungi:
 553: sp_human:
 554: sp_invertebrate:
 555: sp_mammal:
 556: sp_mhc:
 557: sp_organelle:
 558: sp_phage:
 559: sp_plant:
 560: sp_rabbit:
 561: sp_virus:
 562: sp_vertebrate:
 563: sp_unclassified:
 564: sp_bacteriophage:
 565: sp_bacteria:
 566: sp_fungi:
 567: sp_human:
 568: sp_invertebrate:
 569: sp_mammal:
 570: sp_mhc:
 571: sp_organelle:
 572: sp_phage:
 573: sp_plant:
 574: sp_rabbit:
 575: sp_virus:
 576: sp_vertebrate:
 577: sp_unclassified:
 578: sp_bacteriophage:
 579: sp_bacteria:
 580: sp_fungi:
 581: sp_human:
 582: sp_invertebrate:
 583: sp_mammal:
 584: sp_mhc:
 585: sp_organelle:
 586: sp_phage:
 587: sp_plant:
 588: sp_rabbit:
 589: sp_virus:
 590: sp_vertebrate:
 591: sp_unclassified:
 592: sp_bacteriophage:
 593: sp_bacteria:
 594: sp_fungi:
 595: sp_human:
 596: sp_invertebrate:
 597: sp_mammal:
 598: sp_mhc:
 599: sp_organelle:
 600: sp_phage:
 601: sp_plant:
 602: sp_rabbit:
 603: sp_virus:
 604: sp_vertebrate:
 605: sp_unclassified:
 606: sp_bacteriophage:
 607: sp_bacteria:
 608: sp_fungi:
 609: sp_human:
 610: sp_invertebrate:
 611: sp_mammal:
 612: sp_mhc:
 613: sp_organelle:
 614: sp_phage:
 615: sp_plant:
 616: sp_rabbit:
 617: sp_virus:
 618: sp_vertebrate:
 619: sp_unclassified:
 620: sp_bacteriophage:
 621: sp_bacteria:
 622: sp_fungi:
 623: sp_human:
 624: sp_invertebrate:
 625: sp_mammal:
 626: sp_mhc:
 627: sp_organelle:
 628: sp_phage:
 629: sp_plant:
 630: sp_rabbit:
 631: sp_virus:
 632: sp_vertebrate:
 633: sp_unclassified:
 634: sp_bacteriophage:
 635: sp_bacteria:
 636: sp_fungi:
 637: sp_human:
 638: sp_invertebrate:
 639: sp_mammal:
 640: sp_mhc:
 641: sp_organelle:
 642: sp_phage:
 643: sp_plant:
 644: sp_rabbit:
 645: sp_virus:
 646: sp_vertebrate:
 647: sp_unclassified:
 648: sp_bacteriophage:
 649: sp_bacteria:
 650: sp_fungi:
 651: sp_human:
 652: sp_invertebrate:
 653: sp_mammal:
 654: sp_mhc:
 655: sp_organelle:
 656: sp_phage:
 657: sp_plant:
 658: sp_rabbit:
 659: sp_virus:
 660: sp_vertebrate:
 661: sp_unclassified:
 662: sp_bacteriophage:
 663: sp_bacteria:
 664: sp_fungi:
 665: sp_human:
 666: sp_invertebrate:
 667: sp_mammal:
 668: sp_mhc:
 669: sp_organelle:
 670: sp_phage:
 671: sp_plant:
 672: sp_rabbit:
 673: sp_virus:
 674: sp_vertebrate:
 675: sp_unclassified:
 676: sp_bacteriophage:
 677: sp_bacteria:
 678: sp_fungi:
 679: sp_human:
 680: sp_invertebrate:
 681: sp_mammal:
 682: sp_mhc:
 683: sp_organelle:
 684: sp_phage:
 685: sp_plant:
 686: sp_rabbit:
 687: sp_virus:
 688: sp_vertebrate:
 689: sp_unclassified:
 690: sp_bacteriophage:
 691: sp_bacteria:
 692: sp_fungi:
 693: sp_human:
 694: sp_invertebrate:
 695: sp_mammal:
 696: sp_mhc:
 697: sp_organelle:
 698: sp_phage:
 699: sp_plant:
 700: sp_rabbit:
 701: sp_virus:
 702: sp_vertebrate:
 703: sp_unclassified:
 704: sp_bacteriophage:
 705: sp_bacteria:
 706: sp_fungi:
 707: sp_human:
 708: sp_invertebrate:
 709: sp_mammal:
 710: sp_mhc:
 711: sp_organelle:
 712: sp_phage:
 713: sp_plant:
 714: sp_rabbit:
 715: sp_virus:
 716: sp_vertebrate:
 717: sp_unclassified:
 718: sp_bacteriophage:
 719: sp_bacteria:
 720: sp_fungi:
 721: sp_human:
 722: sp_invertebrate:
 723: sp_mammal:
 724: sp_mhc:
 725: sp_organelle:
 726: sp_phage:
 727: sp_plant:
 728: sp_rabbit:
 729: sp_virus:
 730: sp_vertebrate:
 731: sp_unclassified:
 732: sp_bacteriophage:
 733: sp_bacteria:
 734: sp_fungi:
 735: sp_human:
 736: sp_invertebrate:
 737: sp_mammal:
 738: sp_mhc:
 739: sp_organelle:
 740: sp_phage:
 741: sp_plant:
 742: sp_rabbit:
 743: sp_virus:
 744: sp_vertebrate:
 745: sp_unclassified:
 746: sp_bacteriophage:
 747: sp_bacteria:
 748: sp_fungi:
 749: sp_human:
 750: sp_invertebrate:
 751: sp_mammal:
 752: sp_mhc:
 753: sp_organelle:
 754: sp_phage:
 755: sp_plant:
 756: sp_rabbit:
 757: sp_virus:
 758: sp_vertebrate:
 759: sp_unclassified:
 760: sp_bacteriophage:
 761: sp_bacteria:
 762: sp_fungi:
 763: sp_human:
 764: sp_invertebrate:
 765: sp_mammal:
 766: sp_mhc:
 767: sp_organelle:
 768: sp_phage:
 769: sp_plant:
 770: sp_rabbit:
 771: sp_virus:
 772: sp_vertebrate:
 773: sp_unclassified:
 774: sp_bacteriophage:
 775: sp_bacteria:
 776: sp_fungi:
 777: sp_human:
 778: sp_invertebrate:
 779: sp_mammal:
 780: sp_mhc:
 781: sp_organelle:
 782: sp_phage:
 783: sp_plant:
 784: sp_rabbit:
 785: sp_virus:
 786: sp_vertebrate:
 787: sp_unclassified:
 788: sp_bacteriophage:
 789: sp_bacteria:
 790: sp_fungi:
 791: sp_human:
 792: sp_invertebrate:
 793: sp_mammal:
 794: sp_mhc:
 795: sp_organelle:
 796: sp_phage:
 797: sp_plant:
 798: sp_rabbit:
 799: sp_virus:
 800: sp_vertebrate:
 801: sp_unclassified:
 802: sp_bacteriophage:
 803: sp_bacteria:
 804: sp_fungi:
 805: sp_human:
 806: sp_invertebrate:
 807: sp_mammal:
 808: sp_mhc:
 809: sp_organelle:
 810: sp_phage:
 811: sp_plant:
 812: sp_rabbit:
 813: sp_virus:
 814: sp_vertebrate:
 815: sp_unclassified:
 816: sp_bacteriophage:
 817: sp_bacteria:
 818: sp_fungi:
 819: sp_human:
 820: sp_invertebrate:
 821: sp_mammal:
 822: sp_mhc:
 823: sp_organelle:
 824: sp_phage:
 825: sp_plant:
 826: sp_rabbit:
 827: sp_virus:
 828: sp_vertebrate:
 829: sp_unclassified:
 830: sp_bacteriophage:
 831: sp_bacteria:
 832: sp_fungi:
 833: sp_human:
 834: sp_invertebrate:
 835: sp_mammal:
 836: sp_mhc:
 837: sp_organelle:
 838: sp_phage:
 839: sp_plant:
 840: sp_rabbit:
 841: sp_virus:
 842: sp_vertebrate:
 843: sp_unclassified:
 844: sp_bacteriophage:
 845: sp_bacteria:
 846: sp_fungi:
 847: sp_human:
 848: sp_invertebrate:
 849: sp_mammal:
 850: sp_mhc:
 851: sp_organelle:
 852: sp_phage:
 853: sp_plant:
 854: sp_rabbit:
 855: sp_virus:
 856: sp_vertebrate:
 857: sp_unclassified:
 858: sp_bacteriophage:
 859: sp_bacteria:
 860: sp_fungi:
 861: sp_human:
 862: sp_invertebrate:
 863: sp_mammal:
 864: sp_mhc:
 865: sp_organelle:
 866: sp_phage:
 867: sp_plant:
 868: sp_rabbit:
 869: sp_virus:
 870: sp_vertebrate:
 871: sp_unclassified:
 872: sp_bacteriophage:
 873: sp_bacteria:
 874: sp_fungi:
 875: sp_human:
 876: sp_invertebrate:
 877: sp_mammal:
 878: sp_mhc:
 879: sp_organelle:
 880: sp_phage:
 881: sp_plant:
 882: sp_rabbit:
 883: sp_virus:
 884: sp_vertebrate:
 885: sp_unclassified:
 886: sp_bacteriophage:
 887: sp_bacteria:
 888: sp_fungi:
 889: sp_human:
 890: sp_invertebrate:
 891: sp_mammal:
 892: sp_mhc:
 893: sp_org

Qy	61	PRLTNSAMRAGHILDLKLTRGKNGATAFLESIKFHNPDVYTLVLTGQLQPDVDFSNSFGLME	120	DE HYPOTHETICAL 75_4 KDA PROTEIN (FRAGMENT).
Db	61	PRITNSAMRAGHILDLKLTRGKNGATAFLESIKFHNPDVYTLVLTGQLQPDVDFSNSFGLME	120	OS Mus musculus (Mouse).
OC				OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC				OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN				RN [1]
NCBI_TAXID				NCBI_TAXID=1090;
RA				RA SEQUENCE FROM N_A.
RA				RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL				RL DR EMBL: BC004692; AAH04692-1; -.
DR				DR InterPro: IPR001478; PDZ; 1.
PFam				DR PF00552; PDZ; 1.
SMART				DR SMART: SM00228; PDZ; 1.
PROSITE				DR PROSITE: PS00106; PDZ; 1.
KW				KW Hypothetical protein.
FT				FT NONTER 1 1 MW; AADFE37A7A04513D CRC64;
SEQUENCE				SEQUENCE 662 AA; 75422 MW;
				Query Match 48.3%; Score 2484.5; DB 11; Length 662;
				Best Local Similarity 75.1%; Pred: No. 1..3e-138; Indels 15; Gaps 5;
Matches				Mismatches 89; Indels 15; Gaps 5;
Qy				Qy 82 KNGIAIAFLESIKFHNPDVYTLVLTGQLQPDVDFSNSFGLMETSKTECLAGAIGSLOEELNQ 141
Db				Db 1 KNGIAIAFLESIKFHNPDVYTLVLTGQLQPDVDFSNSFGLMETSKTECLAGAIGSLOEELNQ 60
Qy				Qy 142 ERQKEVYLRRQCOLQKELGLAATRGLHOLEADHMRKREVSAHHRPEVLRKDEISSL 201
Db				Db 61 ERQKEVYLRRQCOLQKELGLAATRGLHOLEADHMRKREVSAHHRPEVLRKDEISSL 120
Qy				Qy 202 SLHYNALQKELAASRSRSQEELYLKQELQRANNNSSCCELEFOQLRTASDOESGD 261
Db				Db 121 SLHYNALQKELAATRQHSLOEELYLQELQRASLSSCREPRESLANASNLPPQ 180
Qy				Qy 262 EELNRKENEKLRSLTSFLAKDILQSLDEARGSROLEYERIHSRLRAYERAAEROEQ 321
Db				Db 181 EELNRKENEKLRSLTSFLAKDILQSLDEARGSROLEYERIHSRLRAYERAAEROEQ 240
Qy				Qy 322 YWEKEQFLLOFOFSKMCQCLYBEKVNALQAOVCELOKERDAYSADSQAPEISQVE 381
Db				Db 241 YWEKEQFLLOQFRKTQVQDCELYKEKMTMQLGOVAELQRDQAYTARDRAQEISQRLVE 300
Qy				Qy 362 KDSLRQQVELFDQVEURTQLRQLQAEPPGVLKQEARTRCPREKQLYRMHAICPRD 441
Db				Db 301 KDLARRRYVEFLQVCEURTQLRQLQAEPPGVLKQEARTRCPREKQLYRMHAICPRD 360
Qy				Qy 442 DSDGSILVSYSTESQQLDSLATSRELVSFRSSSPAPSQSQSLYKRAFDSEBPNFSS 501
Db				Db 361 DSDGSILVSYSTESQQLDSLATSRELVSFRSSSPAPSQSQSLYKRAFDLEDFDPES-- 417
Qy				Qy 502 CLEIPEGDPGALPGKAG--DPLHDYELDTAQLPOLESSPQSPGRDVSSEGVLMRR 559
Db				Db 418 -LSFPPEVLEMRLQGATVDDTDLEFENIDGADSQTDSLOG--SSRLNLVSESSVPVR 475
Qy				Qy 560 RPARRILSQVMTLAFOGDLALLEQVLTGQLQISVIGNLTGFIRHRTPGSAADQMLRPTQIVMD 619
Db				Db 476 RPARKLISQVTLAFOGDLALLEQVLTGQLQISVIGNLTGFIRHRTPGSAADQMLRPTQIVMD 535
Qy				Qy 620 YEASEPLPKAVLDTTLEAVGLLIRRQVNTDGKRLQDLEAKVATSGDSF 679
Db				Db 536 YKPTKPSLRLTENTLEAOVGLLIRRQVNTDGKRLQDLEAKVATSGDSF 595
Qy				Qy 680 YIRVNLAMEGRAKGELQYVHCNEVYDQGKQKLYVSMDF 111
Db				Db 596 YIRVNLAMORGGSDELQHNDLHVTMFQGRSCWMAHHNVPTMDMMP-GTIPNYS 654
RESULT	2			Qy 740 RAQQQLLIQD 751
Q99KFO		PRELIMINARY:	PRT:	Q99KFO ID : : :
AC				AC Q99KFO; ID : : :
DT	01-JUN-2001 (TREMBUREL. 17, Created)			DT 01-JUN-2001 (TREMBUREL. 17, Last sequence update)
DT	01-DEC-2001 (TREMBUREL. 19, Last annotation update)			DT 01-DEC-2001 (TREMBUREL. 19, Last annotation update)
RESULT	3			RESULT 3

Q9BVB5	PRELIMINARY;	PRT;	434 AA.					
AC	Q9BVB5;							
DR	01-JUN-2001 (TREMBrel. 17, Created)							
DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)							
DT	01-DEC-2001 (TREMBrel. 18, Last annotation update)							
DE	HYPOTHETICAL 48.5 KDA PROTEIN.							
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
OX	NCBI_TAXID=9606;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RA	TISSUE-CERVIX, CARCINOMA;							
RA	Strausberg R.;							
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.							
DR	EMBL; BC001326; AAH01326..1;							
KW	HYPOTHETICAL protein.							
SQ	SEQUENCE 434 AA.; 48504 MW; D351699781B6D6EB CRC64;							
Query Match	37.1%; Score 1910; DB 4; Length 434;							
Best Local Similarity	100.0%; Pred. No. 6 105;							
Matches	381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
238	MYSCELELQEQLRATASDQESGDEELNRKKEENKLRSLSLAERDILEQSLEARGS	297						
Qy								
Db	1 MYSCELELQEQLRATASDQESGDEELNRKKEENKLRSLSLAERDILEQSLEARGS	60						
Qy	298 RQELEVTHSLRERVAARERQEQYEEKWEEKTQLQFQSKMACQLYREKVNALQAQCEL	357						
Db	61 RQELEVTHSLRERVAARERQEQYEEKWEEKTQLQFQSKMACQLYREKVNALQAQCEL	120						
Qy	358 RQELEVTHSLRERVAARERQEQYEEKWEEKTQLQFQSKMACQLYREKVNALQAQCEL	417						
Db	121 RQELEVTHSLRERVAARERQEQYEEKWEEKTQLQFQSKMACQLYREKVNALQAQCEL	180						
Qy	418 ARTRCPCPREKQRVLVRMHAICPRDDSDCSLVSSTESQQLSDSATSSRELVDSSFRSSPA	477						
Db	181 ARTRCPCPREKQRVLVRMHAICPRDDSDCSLVSSTESQQLSDSATSSRELVDSSFRSSPA	240						
Qy	478 PPQQSLYKRAVADFGEEPWFSSECCLEPGEFDGALPCAKGDPHDYELLTDADLPOL	537						
Db	241 PPQQSLYKRAVADFGEEPWFSSECCLEPGEFDGALPCAKGDPHDYELLTDADLPOL	300						
Qy	538 SSLOQPVSGRLDVSESGYLMRRPARRLSQVITMLAFQGDALLEQISVIGGNITGIFIPHR	597						
Db	301 SSLOQPVSGRLDVSESGYLMRRPARRLSQVITMLAFQGDALLEQISVIGGNITGIFIPHR	360						
Qy	598 YTPGSAADQMALRPGTQIVMF	618						
Db	361 YTPGSAADQMALRPGTQIVMF	381						
RESULT	4							
Q9BXL7	PRELIMINARY;	PRT;	1147 AA.					
ID	Q9BXL7							
AC	Q9BXL7;							
DR	01-JUN-2001 (TREMBrel. 17, Created)							
DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)							
DE	CASPASE RECRUITMENT DOMAIN PROTEIN 11.							
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
OX	NCBI_TAXID=9606;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RX	MEDLINE-21192234; PubMed=11278692;							
RA	Srinivasula S.M., Merriam S., Di Stefano P.S., Alnemri E.S.;							
RA	*CARD11 and CARD14 Are Novel Caspase Recruitment Domain (MACUK) Family Members (CARD)/membrane-associated Guanylate Kinase (MACUK) Family Members							
RT	that Interact with BCL10 and Activate NF-kappaB.";							
RL	J. Biol. Chem. 276:1187-1188 (2001).							
DR	EMBL; AF22641; AAC53402..1; -.							
DR	InterPro; IPR001478; PDZ.							
DR	Pfam; PF00595; PDZ; 1.							
DR	SMART; SM00228; PDZ; 1.							
SQ	SEQUENCE 1147 AA.; 132641 MW; 913A4B015D2B36CC CRC64;							
Query Match	24.1%; Score 1239.5; DB 4;							
Best Local Similarity	30.5%; Pred. No. 7.9e-65;							
Matches	354; Conservative 204; Mismatches 402; Indels 201; Gaps 33;							
Qy	15 DEETLWEMMEHSHRIRIVRCIPSRLTPYLRQAKVLCQDDEEVHSPLRTNSAMRAGHLL	74						
Db	11 EEDALWENVECNRHMLSRVNPALKYPLRQCKVDEQDDEVLNAPMPLPSKTMRAGRLL	70						
Qy	75 DLKTRIGKNGIAFLESLKFHNPDVYTLYTGLQPDVDFSNESGMETSKLTCLELAGAIGS	134						
Db	71 DLHTKGQGYVVPLESFLEYPLKLVGKEPTRFSTIVVEGHEGLTHFLMNEVIK	130						
Qy	135 LQEELNQERGQKEVLLRRCQQL-EHLGLAETRAEGHOLEADISRMKKEVSAHFHEVLR	193						
Db	131 LQOQMKAQDQLCERCELLARQLEDEKONTLTRLTR-LLTQFQERYMKKEFDSNDELYK	189						
Qy	194 LKDEMLSLSLHYSNALQEKELAASRCSLSQEEYLKLQELQVANMVSCELEBOQSLRT	253						
Db	190 VKDDNNYLNAMRYAQLESEEVMAYNRSRDQLETDQLKHRLNK--MEEECKLE-RNQSLKL	246						
Qy	254 ASDQES--GDEELNRKKEEENKLRSLSFSL-----AERDILEQSLDEAR	295						
Db	247 KNDLENRPKEQVLEERENMLTKQNBLQSI1QAGKRSLPLSDKAILDILEHDKREAL	306						
Qy	296 GSROBLVERIHSRERVAARERQEQYKEKEQYKECQLLQFQSKMACQLYREKVNALQAQYC	355						
Db	307 EDROBLVNRVLYNQLEBARAEELKDYLBEKEDBLKCSTLGKDQVYKHMNTVMLQLE	366						
Qy	356 ELQKERDQAYSARDSAQREIISOSLVEKDSLRLRQVFELTDQ-----VCELR	401						
Db	367 EVERDQAFHSRDEATOYSOCUJEKDYKQFRELKEENDENRIMYRREACTIVLNLS	426						
Qy	402 QLROLQAE-----PPGVLKQ-----EARTRECPREK-----428							
Db	427 KLRRSLKDSDNNLDOSLPRNLPTIISQDFGASPRTNQFEADSSTSSEESPSDKYFLPY	486						
Qy	429 429 -----APPSSQSLYKRAVADFGEEPWFSFSCLEPGEFDGALPCAKGDPHDYELLTDADLPOL							
Db	437 EVERLVRHAI-CPRDDSDCSLVSSETEQOLL-----SLSATSSRELVDSSFRSSPA							
Qy	487 HPPORMRNKGIQLORAKSPISLRTSDQAKGHBEGTDDAPSQCSGSLPITSNSTPKMQP	546						
Db	494 ILSOVTMLAFQDALEQIVSGNLTGIFIPHR	597						
Qy	547 PRRSSSIMSTAEPGNGDSVRYKEDAPHR----STVE-EDNDGGFDALDDDSHE	600						
Qy	524 DY-----ELLDTADPOLESSLQPV-----PGRLDVSESGVLMRRPARR	564						
Db	601 RYSGFPSSIHSSSSQSGSLDAYLEQVNLMFKFSLERPFRPSVTSVGHVRGPGPS	-- 658						
Qy	565 ILSOVTMLAFQDALEQIVSGNLTGIFIPHRGTLQVDMRGTQDLEAKVATSGDSFYIRVN	624						
Db	659 --VQHTLNLGDSLTSQTLQGNGARGSFVHSYKPGSLAEKGLREGHQLLLEGCIRG	714						
Qy	625 PLFKAVLDEDTLEAVGLLRRDGFCLSVKVNNTDGKRLQDLEAKVATSGDSFYIRVN	684						
Db	715 ERQSVPLDPCTKEAHWTIQRKVGEGYRKLVKDMDGLTSGDFYIRLN	774						
Qy	685 LAMEGRAKG-ELQVHCNEVLYHVTDMFQGGCWIAHVRNSYTMKDAAHTGIPNYSRAQQ	743						
Db	775 LNISQSLDQACTMSLKDDVYHVRTMYDRHEWPACARVDPFTDILDM-GTIPSYRAQQ	833						
Qy	744 QLIAJQDMTOQ-----CTVTRK-----SSGGPQ-----KLV	771						
Db	834 LLLVKLQLMRHGSREVDGTHTLRALNLTQPSBEALSSTDPRVSPRLSRASFLFGQLL	893						

"CARD9 is a novel caspase recruitment domain-containing protein that interacts with BCL10/CLAP and activates NF-kappa B.";
 RT J. Biol. Chem. 275:41082-41086 (2000).
 RL EMBL: AF311288; AAC28790; 1; -.
 DR InterPro; IPR01315; CARD.
 DR PROSITE; PS50209; CARD; 1.
 SQ SEQUENCE 536 AA; 62631 MW; 6F33089CB7E6BAC9 CRC64;

Query Match 10.0%; Score 512.5; DB: 11; Length 536;
 Best Local Similarity 28.0%; Pred. No. 2.2e-21;
 Matches 162; Conservative 94; Mismatches 215; Indels 91; Gaps 15;

Query 15 DEETLWEMMEHHRRHRRICPSPRLTPYLRQAKVLQCOLDEEYLHSPLTNSAMRAGHLL 74
 Db 6 NDDCWSALESFRVKLISVIDPSRITPLRQCKVLPNPDPEEVLSDFPNLVTKRVGVLL 65

Query 75 DLILKTRGNGATAFLSLELKHFNPDVYTLYTGQPD-----ME 120
 Db 66 DILQRTGHKGYYAVAFLESLEYPPQYKTYGKEPARVFMSMIDASGESGLTQLLMTEVMK 125

Query 121 TSLITECLAGAIGSLOQELQKELAASRCRSLOQELLYLKLQELQRANMVS 240
 Db 126 LQKVQDITALLSS-KDDFKEKRVDSLLRKHQERVQ-----RL 164

Query 181 KREVAHPEVLRKDEMILSLSHYSNALQEKELAASRCRSLOQELLYLKLQELQRANMVS 240
 Db 165 KECECAELSKRKCDENYDMLAHLSEERGAALMNRDQLEVDPOLRSHIMKAE-D 222

Query 241 SCELELQBQ-SLRTASDQESGDEELNLKEENE---KLRSLTFSLAEK-----DI 286
 Db 223 DCYVERKHTKLKRAMEQPSQELLWLOQEKAJJQARQELEASVQEKGKLHRNSPYIQV 282

Query 287 LEOSLDEARGSRQELVERIHSRLRRAAAEROREQYWEEKETQTLQFQRSKMACOLYREK 346
 Db 283 LEEDWQRQALQHQDQANTFSLRKLDRQAEALTRCMBEKEFQCLIAKDMYKDR 342

Query 347 VNALQAQCYCELOKERDQAYSARDSAQREIJSQSVLEKDSLRLRQYFELTDQVCCELTQRLQ 406
 Db 343 IEATLQOMEVEATRQDQATREELQKQVRELDKADELQLQFQT 402

Query 407 QAEPGPVULKQEARTRREPKEPKRVLV--RMHAICPRDDSDCSLVSSTESQL-LSDLSAT 462
 Db 403 ESR--LLAAEGRKQ---QOLDMLLSSDLEDSSPRNSQELSLPQDLEEDAQLSDKGVL 456

Query 463 LDVSFRSSPAPPSQSOSLYKRYAEDFG-----EEWWS 498
 Db 457 ADRESPEQPFVVLINKKHLSQTHDVPSSSEPEKERRRKESFENYRRKRALRMQNSWR 516

Query 499 FSSCLEIPEGDPGALPGAKGD 520
 Db 517 QG-----EGDHGNTTGSDNTD 532

RESULT 7
 ID Q9H257 PRELIMINARY; PRT; 536 AA.

AC Q9H257_01-MAR-2001 (TREMBlrel. 16, Created)
 AC Q9H257_01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 AC Q9H257_01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE CASPASE RECRUITMENT DOMAIN PROTEIN 9.
 OS Homo sapiens (Human).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1] SEQUENCE FROM N.A.
 RX MEDLINE=20576268; PubMed=11053425;
 RA Berlin J., Guo Y., Wang L., Srinivasula S.M., Jacobson M.D.,
 RA Poyet J.-L., Merriam S., Du M.Q., Dyer M.J.S., Robison K.E.,
 RA Di Stefano P.S., Alnemri E.S.;
 RA "CARD9 is a novel caspase recruitment domain-containing protein that

RT interacts with BCL10/CLAP and activates NF-kappa B.";
 RL J. Biol. Chem. 275:41082-41086 (2000).
 DR EMBL: AF311287; AAC28790; 1; -.
 DR InterPro; IPR01315; CARD.
 DR PROSITE; PS50209; CARD; 1.
 SQ SEQUENCE 536 AA; 62267 MW; 6EB1835315B83DES CRC64;

Query Match 9.6%; Score 493; DB 4; Length 536;
 Best Local Similarity 29.5%; Pred. No. 3.2e-21;
 Matches 155; Conservative 92; Mismatches 200; Indels 78; Gaps 14;

Query 15 DEETLWEMMEHHRRHRRICPSPRLTPYLRQAKVLQCOLDEEYLHSPLTNSAMRAGHLL 74
 Db 6 NDDCWSALESFRVKLISVIDPSRITPLRQCKVLPNPDPEEVLSDFPNLVTKRVGVLL 65

Query 75 DLILKTRGNGATAFLSLELKHFNPDVYTLYTGQPD-----ME 120
 Db 66 DILQRTGHKGYYAVAFLESLEYPPQYKTYGKEPARVFMSMIDASGESGLTQLLMTEVMK 125

Query 121 TSLITECLAGAIGSLOQELQKELAASRCRSLOQELLYLKLQELQRANMVS 240
 Db 126 LQKVQDITALLSS-KDDFKEKRVDSLLRKHQERVQ-----RL 164

Query 181 KREVAHPEVLRKDEMILSLSHYSNALQEKELAASRCRSLOQELLYLKLQELQRANMVS 240
 Db 165 KECECAELSKRKCDENYDMLAHLSEERGAALMNRDQLEVDPOLRSHIMKAE-D 222

Query 241 SCELELQBQ-SLRTASDQESGDEELNLKEENE---KLRSLTFSLAEK-----DI 286
 Db 223 DCYVERKHTKLKRAMEQPSQELLWLOQEKAJJQARQELEASVQEKGKLHRNSPYIQV 282

Query 287 LEOSLDEARGSRQELVERIHSRLRRAAAEROREQYWEEKETQTLQFQRSKMACOLYREK 346
 Db 283 LEEDWQRQALQHQDQANTFSLRKLDRQAEALTRCMBEKEFQCLIAKDMYKDR 342

Query 347 VNALQAQCYCELOKERDQAYSARDSAQREIJSQSVLEKDSLRLRQYFELTDQVCCELTQRLQ 406
 Db 343 IEATLQOMEVEATRQDQATREELQKQVRELDKADELQLQFQT 402

Query 407 QAEPGPVULKQEARTRREPKEPKRVLV--RMHAICPRDDSDCSLVSSTESQL-LSDLSAT 462
 Db 403 ESR--LLAAEGRKQ---QOLDMLLSSDLEDSSPRNSQELSLPQDLEEDAQLSDKGVL 456

Query 463 LDVSFRSSPAPPSQSOSLYKRYAEDFG-----EEWWS 498
 Db 457 ADRESPEQPFVVLINKKHLSQTHDVPSSSEPEKERRRKESFENYRRKRALRMQNSWR 516

RESULT 8
 ID Q9H54 PRELIMINARY; PRT; 366 AA.

AC Q9H54_01-MAR-2001 (TREMBlrel. 16, Created)
 AC Q9H54_01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 AC Q9H54_01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE CDNA FLJ13939 FIS, CLONE Y79A1000827.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN [1] SEQUENCE FROM N.A.

RP Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M., Waga S., Hosoi T., Kaku Y., Koda Ira H., Ono Y., Takaguchi S., Takahashi M., Chiba S., Murakawa K., Ono Y., Saito K., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamanoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Nonomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG 2000) to the EMBL/GenBank/DDJB databases.
 DR EMBL; AK024001; BAB14766; 1; -.

DR	InterPro; IPRO001315; CARD.	DR	PROSITE; PS50209; CARD; 1.	DR	PROSITE; PS50002; SH3; 1.	DR	PROSITE; PS50002; SH3; 1.		
SO	SEQUENCE 366 AA; 42980 MW;	SO	SEQUENCE 366 AA; 42980 MW;	SO	SEQUENCE 366 AA; 42980 MW;	SO	SEQUENCE 366 AA; 42980 MW;		
Query Match	Score 7.7%; Best Local Similarity 30.7%; Matches 116; Conservative	Score 396.5%; Pred. No. 9.3e-16; Mismatches 65; Matches 146;	Length 366; Indels 51; Gaps 8;	Score 6.8%; Best Local Similarity 20.9%; Matches 240; Conservative	Score 351.5%; Pred. No. 3.3e-12; Mismatches 376; Indels 363; Gaps 49;	Length 1695; Score 19.878 MW; Pred. No. AE0E7D09007316F1 CRC64;	Length 1695; Score 19.878 MW; Pred. No. AE0E7D09007316F1 CRC64;		
QY	15 DEETLWEMMESHRRHIVRCICPSRLTPYLQARVKVLCOLDEEVTLHSPLRNTSAMRAGHLL 74	QY	68 MRAGHLLDLIKTRGRNGAIAFLESLKRFHNPDVVTL-VTGLQDPDVDF-----SNTSG 117	QY	68 MRAGHLLDLIKTRGRNGAIAFLESLKRFHNPDVVTL-VTGLQDPDVDF-----SNTSG 117	QY	68 MRAGHLLDLIKTRGRNGAIAFLESLKRFHNPDVVTL-VTGLQDPDVDF-----SNTSG 117		
Db	6 NDDCWNVJEGFRVLTTSVYDPSRITPYLRQCKVLNPDPDEEVQLSDPBNVIRKRYGVLL 65	Db	67 LKVGDILISVNEYNDG-----RSHDAVEALKAGMEARNEIKRPSINPPKKVNDN 118	Db	67 LKVGDILISVNEYNDG-----RSHDAVEALKAGMEARNEIKRPSINPPKKVNDN 118	Db	67 LKVGDILISVNEYNDG-----RSHDAVEALKAGMEARNEIKRPSINPPKKVNDN 118		
QY	75 DLLATRGRNGAIAFLESLKRFHNPDVVTLYTGLQDP-----VDFSNFSGL-----ME 120	QY	118 LMET----SKLTECLAGAIGSLQEELNQEQKVEYLRLRQOLQHGAETRAEGLHLAEQHGLAETRAEGLHL 173	QY	118 LMET----SKLTECLAGAIGSLQEELNQEQKVEYLRLRQOLQHGAETRAEGLHL 173	QY	118 LMET----SKLTECLAGAIGSLQEELNQEQKVEYLRLRQOLQHGAETRAEGLHL 173		
Db	66 DILORTGHGYVAPLESLELYPOLYKVTGKPARVFSMIDASGSGLTQOLLTEVMK 125	Db	119 LNDSKVNGSKTDSERGRSKKPKMENEQESOKKH-----HSSHQK 158	Db	119 LNDSKVNGSKTDSERGRSKKPKMENEQESOKKH-----HSSHQK 158	Db	119 LNDSKVNGSKTDSERGRSKKPKMENEQESOKKH-----HSSHQK 158		
QY	121 TSKLTCLAGAIGSLQEELNQEQKVEYLRLRQOLQHGAETRAEGLHLADHSRM 180	QY	174 EADHSRMKREVSAYSAHFEVLRKDEMILSLSHYSNALQEKELAASCRCSRSLQEEYLKLQKQEL 233	QY	174 EADHSRMKREVSAYSAHFEVLRKDEMILSLSHYSNALQEKELAASCRCSRSLQEEYLKLQKQEL 233	QY	174 EADHSRMKREVSAYSAHFEVLRKDEMILSLSHYSNALQEKELAASCRCSRSLQEEYLKLQKQEL 233		
Db	126 LQKQVQDTALLSS-KDFTKELRKVDSSLRKHQERY-----RL 164	Db	159 EDDHDSEKPR-----HRSKREKGEGEKSRS-----186	Db	159 EDDHDSEKPR-----HRSKREKGEGEKSRS-----186	Db	159 EDDHDSEKPR-----HRSKREKGEGEKSRS-----186		
QY	181 KREVSAYSAHFEVLRKDEMILSLSHYSNALQEKELAASCRCSRSLQEEYLKLQKQELQRANMV 240	QY	234 QANMVSSCSELBLQEOLSRTA-----DQESGDE-----ELNRNKEENENKLRSL 277	QY	234 QANMVSSCSELBLQEOLSRTA-----DQESGDE-----ELNRNKEENENKLRSL 277	QY	234 QANMVSSCSELBLQEOLSRTA-----DQESGDE-----ELNRNKEENENKLRSL 277		
Db	165 KEECAGSRSPLRKCRKEENYDMLMLAHQSEEKGAAALMNRNDRDQLEIDOLKHSEMKAED 222	Db	187 HNKSLLEENDADTDRSENSHRSRSKRNREDNAGEKSRSRSK 246	Db	187 HNKSLLEENDADTDRSENSHRSRSKRNREDNAGEKSRSRSK 246	Db	187 HNKSLLEENDADTDRSENSHRSRSKRNREDNAGEKSRSRSK 246		
QY	241 SCELLEQEQ-SLRTASDQESBESQELNURKEKE-----NEKLRSLTFSLAKK-----DI 286	QY	278 -----TFSLABKDILEDARGSRQELVERIHSRLR-----AVAAERREQYWE 324	QY	278 -----TFSLABKDILEDARGSRQELVERIHSRLR-----AVAAERREQYWE 324	QY	278 -----TFSLABKDILEDARGSRQELVERIHSRLR-----AVAAERREQYWE 324		
Db	223 DCKVERKHTLKRNAMEQRPSQELWEOQEKKALQARVOQEELASVQEKGKLRSPYQV 282	QY	325 -----EKEQTILQFOFSKMACQLYR-----EKVNALO-----AQVSARDSA 371	QY	325 -----EKEQTILQFOFSKMACQLYR-----EKVNALO-----AQVSARDSA 371	QY	325 -----EKEQTILQFOFSKMACQLYR-----EKVNALO-----AQVSARDSA 371		
QY	287 LEQSLDEARGSRQELVERIHSRLRERVAAREQREQWEEKTQTLQFOFSKMACQLYREK 346	Db	307 DPTEEKSLR-SKSKTREDLSKTTSESEKPKSOSKASLGGMEMENSDRQSKTREDA 365	Db	307 DPTEEKSLR-SKSKTREDLSKTTSESEKPKSOSKASLGGMEMENSDRQSKTREDA 365	Db	307 DPTEEKSLR-SKSKTREDLSKTTSESEKPKSOSKASLGGMEMENSDRQSKTREDA 365		
Db	283 LEEDRQALRQHQEQANTIFSLRKDLRQEARRLRCMEEKEMFELQCLALRKDSKMYKDR 342	QY	372 QREISQSLIVEKDSLRLRQYFELTDQCELTORLQQAEPGVYKOFARTREPCCPREKORL 431	QY	372 QREISQSLIVEKDSLRLRQYFELTDQCELTORLQQAEPGVYKOFARTREPCCPREKORL 431	QY	372 QREISQSLIVEKDSLRLRQYFELTDQCELTORLQQAEPGVYKOFARTREPCCPREKORL 431		
QY	347 VNALQAQVCEQKERDQA 364	Db	366 GEPEKARMSTOS-KPNIESDK-HRSRSKIRE--DTEGTEKPRSRSTORKPNIDMEKL 420	Db	366 GEPEKARMSTOS-KPNIESDK-HRSRSKIRE--DTEGTEKPRSRSTORKPNIDMEKL 420	Db	366 GEPEKARMSTOS-KPNIESDK-HRSRSKIRE--DTEGTEKPRSRSTORKPNIDMEKL 420		
Db	343 TEAILLQMEEVAIERDQS 360	QY	432 VRMHAICRDDSDCSLYSTESQOLSDSATSSRELVDSSRFSSTAPPSSQOSLYKRYAED 491	QY	432 VRMHAICRDDSDCSLYSTESQOLSDSATSSRELVDSSRFSSTAPPSSQOSLYKRYAED 491	QY	432 VRMHAICRDDSDCSLYSTESQOLSDSATSSRELVDSSRFSSTAPPSSQOSLYKRYAED 491		
Db	595 HKEDQKSSKITPPDRHSSPVPSNDNHDTKDSWAATPKKHFSSAAHNPQEVEDMNAKI 654	Db	421 NSSRA--NEDGEKASSRSRSRSDLRTSKSKESE-----HRSRSKEALNVDTI--KVVLN 474	Db	421 NSSRA--NEDGEKASSRSRSRSDLRTSKSKESE-----HRSRSKEALNVDTI--KVVLN 474	Db	421 NSSRA--NEDGEKASSRSRSRSDLRTSKSKESE-----HRSRSKEALNVDTI--KVVLN 474		
RESULT	9	Q9BKL2	PRELIMINARY;	PRT;	1695 AA.	QY	556 --LMRRPARRILSQV-----TMIAFQGDLALLEQISVGUNLGTGIFTTHRTPGSAADQNLRL 610	QY	556 --LMRRPARRILSQV-----TMIAFQGDLALLEQISVGUNLGTGIFTTHRTPGSAADQNLRL 610
AC	Q9BKL2	AC	Q9BKL2	AC	FGEFPWSSCSCLETPGGDGA-LPGAKA-----GDPHDYELLDTADU-----533	Db	475 RGDQGYGSSLGQQIFVKBDLAKDSPAAKAQNLEKGDITYREINGTPLDNKISPECIELIRGA 534	Db	475 RGDQGYGSSLGQQIFVKBDLAKDSPAAKAQNLEKGDITYREINGTPLDNKISPECIELIRGA 534
DT	01-JUN-2001 (TRIMBLrel. 17, Last sequence update)	DT	01-JUN-2001 (TRIMBLrel. 17, Last sequence update)	QY	534 -----POLESSQVPSVPSGRDLSSE-----SGV-----555	QY	534 -----POLESSQVPSVPSGRDLSSE-----SGV-----555	QY	534 -----POLESSQVPSVPSGRDLSSE-----SGV-----555
DT	01-DEC-2001 (TRIMBLrel. 19, Last annotation update)	DE	TIGHT JUNCTION PROTEIN ZO-1.	Db	535 SETLITLTYKKPKLEQDVYPPEKFKSLEVNDERPKSKPESSGKQKQEVYSEPSLSKTKEQKMKV 594	Db	535 SETLITLTYKKPKLEQDVYPPEKFKSLEVNDERPKSKPESSGKQKQEVYSEPSLSKTKEQKMKV 594	Db	535 SETLITLTYKKPKLEQDVYPPEKFKSLEVNDERPKSKPESSGKQKQEVYSEPSLSKTKEQKMKV 594
GN	ZO-1.	GN	Hydra attenuata (Hydra vulgaris).	QY	556 -----	QY	556 -----	QY	556 -----
OS	Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;	OC	Hydridae; Hydria; NCBI_TaxID=6087;	Db	595 HKEDQKSSKITPPDRHSSPVPSNDNHDTKDSWAATPKKHFSSAAHNPQEVEDMNAKI 654	Db	595 HKEDQKSSKITPPDRHSSPVPSNDNHDTKDSWAATPKKHFSSAAHNPQEVEDMNAKI 654	Db	595 HKEDQKSSKITPPDRHSSPVPSNDNHDTKDSWAATPKKHFSSAAHNPQEVEDMNAKI 654
RN	[1]	SEQUENCE FROM N.A.	RP	RP;	1695 AA.	QY	556 --LMBRQFQGDLALLEQISVGUNLGTGIFTTHRTPGSAADQNLRL 610	QY	556 --LMBRQFQGDLALLEQISVGUNLGTGIFTTHRTPGSAADQNLRL 610
XR	MEDLINE-21025768; PubMed=11225567;	RA	K. Yan L., Zhang J., Sarris M. P. Jr.;	RA	AF230482; AAK28322; 1.	Db	655 ERKSNRKLERRMQLSPNNAKKEGLK 713	Db	655 ERKSNRKLERRMQLSPNNAKKEGLK 713
RT	"Molecular and biological characterization of a zonula occludens-1 homologue in Hydra vulgaris, named HZO-1.";	RT	De Dev, Genes Evol. 2000; 14:61-66(2000).	RT	IPRO00906; J05.	Db	714 PGDQJIMCNEDF-----ENITREEAVILLALPDDSVLWVKQSTFDQ1IKE 762	Db	714 PGDQJIMCNEDF-----ENITREEAVILLALPDDSVLWVKQSTFDQ1IKE 762
DR	HSSP; P29476; 10QAV.	DR	IPRO00619; Guanylate_kin.	DR	IPRO001478; PD2.	QY	668 LEAKYATSGDSFYIRVNLYAMEGRK-GELQVHCMBVLYHTDMFQG-CGGCHWHAHV--NS 723	QY	668 LEAKYATSGDSFYIRVNLYAMEGRK-GELQVHCMBVLYHTDMFQG-CGGCHWHAHV--NS 723
DR	InterPro; IPRO00619; Guanylate_kin.	DR	InterPro; IPRO001478; PD2.	DR	IPRO001452; SH3.	Db	763 L-----GDNFFIRVNFDHAEKANTNELTFRGEIFNYRDTMVGGLIGWQQRVGRNA 815	Db	763 L-----GDNFFIRVNFDHAEKANTNELTFRGEIFNYRDTMVGGLIGWQQRVGRNA 815
DR	SMART; SM00072; GUIC.	DR	InterPro; IPRO00906; J05.	DR	Pfam; PF00625; Guanylate_kin.	QY	724 YTMKDTAHTCTIPNSRAQOQLIAQDNTQOCITV-----RKPSSGPQKLVRISM 776	QY	724 YTMKDTAHTCTIPNSRAQOQLIAQDNTQOCITV-----RKPSSGPQKLVRISM 776
DR	SMART; SM00329; SH3; 1.	DR	Pfam; PF00595; PD2.	DR	Pfam; PF00791; P05; 1.	Db	816 QMLE---REGLPNPSRNEO---LATAQKMEERQTLTPSKSLRRNSIGGTLKKQKFQS 869	Db	816 QMLE---REGLPNPSRNEO---LATAQKMEERQTLTPSKSLRRNSIGGTLKKQKFQS 869
DR	SMART; SM00228; PD2; 3.	DR	SMART; SM00228; PD2; 3.	DR	DKAKASPLRLSFLRSGDQGOLDPSRMEGSTCFRAESCLTLVPTLYMPHRPARPVPV-----LVP 834	QY	777 DKAKASPLRLSFLRSGDQGOLDPSRMEGSTCFRAESCLTLVPTLYMPHRPARPVPV-----LVP 834	QY	777 DKAKASPLRLSFLRSGDQGOLDPSRMEGSTCFRAESCLTLVPTLYMPHRPARPVPV-----LVP 834
DR	SMART; SM00218; Z05; 1.	DR	DR	DR	DR	DR	870 DRLD---ELSFTEGYQIPA-----YERVVKVADP-----NRPVVVLGPJRD 908	DR	870 DRLD---ELSFTEGYQIPA-----YERVVKVADP-----NRPVVVLGPJRD 908
DR	PROSITE; PS50052; GUANYLATE_KINASE-2; 1.	DR	PROSITE; PS50052; GUANYLATE_KINASE-2; 1.	DR	PROSITE; PS50052; GUANYLATE_KINASE-2; 1.	DR	835 RAVGKILSEKLCLOGFFKKCLAEYLSQLQEEYEAWSORGDIQEGEVSGGRCWTRHAYESL 894	DR	835 RAVGKILSEKLCLOGFFKKCLAEYLSQLQEEYEAWSORGDIQEGEVSGGRCWTRHAYESL 894

Qy	471 PRSSPAPPSSQSLKRAEDEFGE-----EPWSFSSCLEIPEGDPGALPGAKAGD	520	Qy	373 REISQSILVEKDSLRRQWFEILDQVCEIRTQLRQAEPPGVLKQEARTRECPREKQRLY	432
Db	487 LRESQEGREVQREGEAKERLRLSGEAKERLRLSNEELSAVKAESERSLSKLANE-D	545	Db	365 KEISSK-AREOLNANCYSL-QEGDITRHTNCGDTMSIKEAKUIIDGC-KERUNLV	419
Qy	521 PHDYELLDIA-----DLPOLESSLQPVSPGRLDVSESGVLMRRPARRILSQVTMIAF	574	Qy	433 RMHAICPRDD-SDCSLVSSTESQQLSDLSAT-----SSRELVDSSRSSSPAPPSQ	481
Db	546 KEQKLALLEARTAVGKEAGELRTGLQEVERSRLER-----RRELQERRQMMLDS	597	Db	420 VLRTDNTAQSOLNNSAHOASGNIYATHPOQYSGCCSSNNNLEDPYLPGG-ASYSS	478
Qy	575 QGDALIEQISITGGNL-----TGIFTIURTPGSAADOMALRGTCIQWMDYE	621	Qy	482 QSLYKRVAEDFGEPPNSFSSCLEIPEGDPGALPGAKADPHDYELLDTADLPOLESSLQ	541
Db	598 ENTRLGRELALQGRALGEAAKEERETTIGLORLLKGASLEV-MROLQVAQRKLO	656	Db	479 ONLYQEPTRTSNGPNINGLN-DEKSNLTPRGRSRGP---IMDGVLSQLD--R	528
Qy	622 ASEPLEKAVLEDDTLEAVAGLILRRYDVGFCCISLKVUNTDGYKRLQDLEAKVATSGDSFY	681	Qy	542 PVSPGRR----LDV-----SESVLM-----RRRPARRILSQTMLAFOGDALL	580
Db	657 EQEGER-----TREERLIGGLEEARGL-----TEKQDLDHARGLELKLEAAEAE-LGL	706	Db	529 PVTPTGRSAAIDEPPRPPPPRGSSGAAQEDFYSSRRLQBEROSEPREFSQEGSV	588
Qy	682 FVNLANEMGRANG-----ELQVHCNEVL-HVYDTMFEOGCMWA-----HRVNSYTM	726	Qy	581 EOISVIGGNLTCIFIFHRVTPGSAADOMALRPTQIWNVDYEASEPLEKAVLEDDTLEAV	640
Db	707 RLS-AEGRAGLAELARVEQVRRAEAERQGLGLRSALRQLGLGRAPSAPRPPVGSPA	765	Db	589 -GIRLTCGNEAGIFVTAQVQGSPASUQGLMPDKILKVN-----DMNDNGVTREBEAV	639
Qy	727 KDTAAHGTTIPYNSRAQQQLALIQLDMMTOCCTVTRPSSGCFQKLIVRVSMDAKASPLRL	786	Qy	641 GLLRRVQFGCCLSVNVNTDGYKRLQDLEAKVATS-GDSFYTRVNLAEMGRAKGEJQVH	698
Db	766 RDAPAGSGE-----GLNSPTELECPGSPQSPGPA-----TSPASP---	803	Db	640 LFL-----LSLQDIDLIVYQCKEYDEVVTNQFGDSFHKTHTCDNPNSKGEMAFK	691
Qy	787 SFDRGOLDPSRMEGSSTCFWAESCCLTVPTLVWPHRPARPVLVPLVRAWKILSEKLC	846	Qy	699 CNEVLYHTDTMFQG-CGCWVAHVNNSYTMKDTAAHGTPYNSRAQQQLALIQLDMMQCT	757
Db	804 -----DLDPEAVRGALREFLQEL-----RQAQRERDEL--RVCQSLANROLA	843	Db	692 AGDVFRVFTDLINGVYVQSWVQVIGRQHE-MQRGTPKNSRAEEL-----APAQEN	742
Qy	847 DLQGFKKKCL-----AEYLQEEWAEWSORGDIQIQEVEGSSGROWVTHAVESTMKNTHALL	903	Qy	758 VTRKPSSGGPQKLVRIVSMSDKAKASPURLSPDQGLDPSR-----MEGSSTCFWAESCL	811
Db	844 EMEAERDSATSRARQLOKAVSEEEARRSV-DGRLSGVQAEEL--ALOEESVRSERER	988	Db	743 ATKK-----ENNANEERGNFERRRSTHRRSKSLSSRNWDYVFSESIS	786
Qy	904 DVQDLSVCTLR-MDIFPIVHVSYNEKMAN-----KLKGKLGRLGTSEQLEBAA---	953	Qy	812 TLVPTYLWPHRPARPVLVPLVRAVKILSEKLCILQGFKKCLAEWLQSQEEYEAWSQRG	871
Db	899 RATLDQVATLERSLSQATESERAS-QEKISMKANETKLEGDKRLL---KEVLDASESR	953	Db	787 KEPAYERVVLRHPGFVPVVLF-GPYSDLARLLA-KDFPDKFESTPLQDDDKSA-	838
Qy	954 -----RQEGDLDRAPCLYSSLAQDGNAEDDGLSLSCYRQIADEQKKWVTE 1000	1000	Qy	872 DIIQEGEVSGGRCWVTRHA-VESLMEKNTAHLQDLSVCTLHRMDIFPIVHVSYNEK	930
Db	954 TVKLELQRSSLEGLORSRL-----GLSDREAOAQALQDRVDSLQROVADSE 1000	1000	Db	839 -----ATSGKCRIVLRSNIRDVMDSKGHALDITPNAVDLBNYAFQPVVFLKTDISK	891
RESULT	14	15	Qy	931 -MAKKLKGQLQRLG-TSEQOLLEAARQEQQGDLRAPGLYSSLAPDGQ-----SDL	978
ID	Q960N4	PRELIMINARY;	PRT;	974 AA.	
AC	Q960N4;				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
NCBI_TaxID	7227 (TREMBLrel. 19, Last annotation update)				
DE	LDA3161P.				
GN					
OS	Drosophila melanogaster (Bruit fly); Insecta;				
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Phydrodriidae; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N_A.				
RC	STRAIN=N_A;				
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,				
RA	Champe M., Chavez C., Dorett V., Frise E., George R.,				
RA	Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mongall C.J.,				
RA	Nunoo J., Paclet J., Paragas V., Park S., Phouanenavong S., Wan K.,				
RA	Yu C., Lewis S.E., Rubin G.M., Celiker S.,				
RA	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL:AY051965; AAK93389.1;				
DR	SEQUENCE 974 AA; 107338 MW;				
DR	5FA3A7D/D0922C2E CRC64;				
SQ					
Query Match	5.58;	Score 282; DB 5; Length 974;			
Best Local Similarity	22.7%;	Pred. No. 2e-08;			
Matches	155; Conservative 111; Mismatches 276; Indels 142; Gaps 30;				

Query Match Score 282; DB 5; Length 974;
 Best Local Similarity 22.7%; Pred. No. 2e-08;
 Matches 155; Conservative 111; Mismatches 276; Indels 142; Gaps 30;

Sequence From N_A.
 PRELIMINARY; PRT; 1367 AA.
 ID Q94880
 AC Q94880;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE TAMA.
 GN PD OR TAMO OR CG9763.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ebhydrodriidae; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N_A.
 RC STRAIN=CAUTION-S;
 RA Takahisa M., Togashi S., Suzuki T., Kobayashi M., Murayama A.,
 RA Kondo K., Miyake T., Ueda R.;
 RA RT "The Drosophila tamou gene, a component of the activating pathway of
 extrametacochætae expression, encodes a protein homologous to
 mammalian cell-cell junction-associated protein ZO-1.";
 RA RT Genes Dev. 10:1783-1795 (1996).
 RL

```

DR EMBL; D83477; BAA11923.1; -
DR HSSP; Q12923; 3P02;
DR FlyBase; FBan000177; pyd.
DR InterPro; IPR000019; Guanylate_kin.
DR InterPro; IPR01478; PDZ.
DR InterPro; IPR00152; SH3.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDE; 3.
DR SMART; SM00072; Gukc; 1.
DR SMART; SM00228; PDZ; 3.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR SEQUENCE AA: 148347 MW: 571C4566C6B6BF8 CRC64;
SQ

```

Query	Match	Score	DB	Length
Qy	REISQSILVERKDSLRRQVELTDOVCELTQRLQARPPGVLKQEARTREPQPREKORLY	5.5%	282	5; Length 1367;
Qy	373 REISQSILVERKDSLRRQVELTDOVCELTQRLQARPPGVLKQEARTREPQPREKORLY	22.7%	22.7	; Pred. No. 3.1e-08;
Db	194 KELTSKK--AREQLNANGYSL--QEGDITRINTNCGDTMSKREAKKTIDGG--KERINLV		111	; Mismatches 2/6; Indels 142; Gaps 30;
Qy	433 RMHATCPRD-SDGSLSVSSTESQLSLSAT-----SSRELVDSEFSSSPAPPSQ	4.81		
Db	249 VLRDTNTQAVSQLINNNSASHASAGNYATHQPVQSGCCSNNNLDPYLGCG-ASYSS		1	
Qy	482 QSLYKRAEDFGEEPFWSFSSCLEIPEGDGAFLPGAKAGDPHLDYELLDTADIPQLESSLQ	5.41		
Db	308 QNLIVQPPTPTSGNPNTNGNLN--DEKSNLTRGRSRGP----IMDGVSQQLD---R	3.57		
Qy	542 PYSPGR-----LDV-----SESGLM-----RRRPARRILSQTMLAFOGDALL	5.80		
Db	358 PVTPTPRGRSAIDEDEPRPPPGRSGGAQEDFYSSRQLYEROSAAPPFRISFQKGCSV	4.17		
Qy	581 EQISWTGGNTGIFTIHRVTPGSSAQMALRPGTQIVMDYEASEPLFKAVLDTTLEAV	6.40		
Db	418 -GIRUTGGNFAGIFYTAOPGSPASLQGIMPDKILKN-----DMDMNCVTREAV	4.68		
Qy	641 GLLRRVDFGCCSVKVNTQGYKLQLDEAKVATS--GDSFYIRVNLAAMEGRAKGELOHV	6.98		
Db	469 LFL-----LSLQDRIDLIVYCKEDEVTTNQRGDSFHKTFCNDNP-SRGEMAFK	5.20		
Qy	699 CNEVLIHTDMFQG-CGCWHIAHRYNSYTMKDTAAHGHTIFNYSRAQQQLIALIDDMTQOCT	7.57		
Db	521 AGDVERVIDIHLNGVYGSNQVLYKIGRQHE-MQRGVLPNKNSRAEEL-----ATAQFN	5.71		
Qy	758 VTRKPSSGGPKQLYRIVSMKDAKASPLRUSFDQLDPSR----MEGSSTCFWAESCL	8.11		
Db	572 ATKK-----EMMANESRGNFERRRSTIRRKSLSRENWDVVFSDSTS	6.15		
Qy	812 TLVPXTLVPHPRPARPVLVPRAVGKILSEKILCQLQFPKCHAEYLQSQEYEAWSORG	8.71		
Db	616 KFPAYERXVLHHPGIVRPVVL-GPVSDLARERLA-KDEPDKSTPLODDDKSA	6.67		
Qy	872 DIIQEGEVSGGRCVYTRHA-WEISLMEKNTNHALLDYOLDSVCTLHRMD1FPTVTHSYNEK	9.30		
Db	668 -----ATSKRKCRVRLSNRDRMGRKIALLDITPNA'DRLNAYAQF'PVVFLRTDSK	7.20		
Qy	931 -MAKKIKKGLOPLG-TSEEQLEEARQEEGDLDRAPCLYSSLAPDGW-----SDL	9.78		
Db	721 HVIKQLRHGLEKAHKSSKKLLEOCOK-----LERV-----WSHFSSTQIALSDE	7.65		
Qy	979 DGLLSCVYRQAIADEQQKVWTEQS 1002			
Db	766 ESWYRKLDSDLQSGAVMSES 789			

Result No.	Score	Query	Match	Length	DB ID	Description
1	5149	100.0	1004	1	CARE_HUMAN	Q9bx16 homo sapien
2	3869.5	75.2	999	1	CARE_MOUSE	Q99kf0 mus musculus
3	1239.5	24.1	1147	1	CARB_HUMAN	Q9bx17 homo sapien
4	887.5	17.2	1032	1	CARA_HUMAN	Q9bw7 homo sapien
5	886.5	17.2	1021	1	CARA_MOUSE	P88660 mus musculus
6	512.5	10.0	536	1	CAR3_RAT	Q9py0 rat/rattus norvegicus
7	493.	9.6	536	1	CAR9_HUMAN	Q9bh257 homo sapien
8	289.	5.6	1745	1	ZOI_MOUSE	P39447 mus musculus
9	280.	5.4	1167	1	ZO2_MOUSE	P320u1 mus musculus
10	270.5	5.3	1736	1	ZOI1_HUMAN	Q07157 homo sapien
11	252.6	5.0	1174	1	ZO2_CANFIA	Q5168 canis familiaris
12	254.5	4.9	1190	1	ZO2_HUMAN	Q16y2 homo sapien
13	249	4.8	4684	1	PLE1_HUMAN	Q15149 homo sapien
14	246.5	4.8	905	1	ZO3_MOUSE	Q9gxy1 mus musculus
15	238.5	4.7	1950	1	MYH9_HUMAN	P35579 homo sapien
16	238.5	4.6	898	1	ZO3_CANFIA	Q62683 canis familiaris
17	236.5	4.6	4473	1	PLE1_CRGCR	Q9j55 cricetus
18	236	4.6	3220	1	CENF_HUMAN	P19454 gallus gallus
19	235	4.6	1959	1	MYH9_CHICK	P14105 rattus norvegicus
20	234.5	4.6	4687	1	PLE1_RAT	P30427 rattus norvegicus
21	226.5	4.4	1976	1	MYHA_RAT	Q91lto rattus norvegicus
22	226	4.4	1972	1	MYHB_RABBIT	P35748 oryctolagus cuniculus
23	225.5	4.4	2349	1	TPR_HUMAN	P22720 homo sapien
24	225	4.4	2704	1	BPA1_HUMAN	P03001 homo sapien
25	223	4.3	1939	1	MYH1_CHICK	P12882 homo sapien
26	219	4.3	1972	1	MYHB_HUMAN	P35749 homo sapien
27	218	4.2	2220	1	GOGA_HUMAN	P13439 homo sapien
28	217.5	4.2	933	1	ZO3_HUMAN	Q95049 homo sapien
29	217.5	4.2	1976	1	MYHA_HUMAN	P35580 homo sapien
30	215	4.2	1072	1	MYHB_MOUSE	Q88638 mus musculus
31	215.5	4.2	1087	1	AKA9_RABBIT	P386628 oryctolagus cuniculus
32	214	4.2	1407	1	TRH2_RABBIT	P37709 oryctolagus cuniculus
33	214	4.2	1941	1	MYH2_HUMAN	Q9ukx2 homo sapien

ALIGNMENTS					
RESULT 1					
CARE_HUMAN					
ID	CARE_HUMAN	STANDARD;	PRT;	1004 AA.	
AC	OBXXL6; Q9BYB5;				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DE	Caspase recruitment domain protein 14 (CARD-containing MAGUK protein				
DE	(Carma 2).				
GN	CARD14 OR CARMA2.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
NCBI_TAXID	9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEIDLINE=21192234; PubMed=11278692;				
RA	Bertin J., Wang L., Guo Y., Jacobson M.D., Poyet J.-L., Srinivasula S.M., Merriam S., Distefano P.S., Alnemri E.S.;				
RT	"CARD11 and CARD14 are novel caspase recruitment domain (CARD)-membrane-associated guanylate kinase (MAGUK) family members that interact with Bcl10 and activate NF-kappaB.";				
RT	J. Biol. Chem. 276:11877-11882(2001).				
RL	[2]				
RN					
RP	SEQUENCE FROM N.A.				
RX	PUBMED=11356195;				
RA	Gaido O., Martinon F., Micheau O., Bonnet D., Thome M., Tschoopp J.;				
RT	"CARD11, a CARD-containing binding partner of Bcl10, induces Bcl10 phosphorylation and NF-kappaB activation.";				
RT	FEBS Lett. 496:121-127(2001).				
RL	[3]				
RN					
RP	ERRATUM.				
RA	Gaido O., Martinon F., Micheau O., Bonnet D., Thome M., Tschoopp J.; FEBS Lett. 505:195-198(2001).				
RL	[4]				
RP	SEQUENCE OF 1-740 FROM N.A.				
RC	TISSUE=cervix, and Colon;				
RT	Strausberg R.;				
RA	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.				
RL	-1 - FUNCTION: Activates NF-kappaB via Bcl10 and IKK. Stimulates the				
CC	-1 - SUBUNIT: CARD14 and Bcl10 bind to each other by CARD-CARD interaction.				
CC	-1 - SUBCELLULAR LOCATION: Cytoplasmic.				
CC	-1 - TISSUE SPECIFICITY: Expressed in placenta. Also detected in HeLa S3 cells, but not in the other cancer cell lines tested.				
CC	-1 - PHOSPHORYLATION OF BCL10.				
CC	-1 - SIMILARITY: CONTAINS 1 CARD DOMAIN.				
CC	-1 - SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.				
CC	-1 - SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.				
CC	-1 - CAUTION: Supposed to contain a SH3 domain which is not detected by PROSITE, Pfam or SMART.				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way				
CC					

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	DR;	EMBL; AF322642;	PAC53403_1;	Y	601 GSADQMAIRPQTGTOIVMVYDYEASEPLEKAVLEDTTLEFAAGLIRRVDGFCCISVVKNTDG 660	
CC	DR;	EMBL; AAK54453_1;	-;	Db	601 GSADQMAIRPQTGTOIVMVYDYEASEPLEKAVLEDTTLEFAAGLIRRVDGFCCISVVKNTDG 660	
CC	DR;	EMBL; EC018142;	PAH18142_1;	Y	661 YKRLLQDLEAKVATSGSDPYIYNLNAMEGRAKEQVHNEVLYHVTDTMFQGGCWHAHR 720	
CC	DR;	EMBL; BC001326;	PAHO1326_1;	Db	661 YKRLLQDLEAKVATSGSDPYIYNLNAMEGRAKEQVHNEVLYHVTDTMFQGGCWHAHR 720	
CC	DR;	InterPro; IPR000019;	Guanylate_Kin.	Y	661 YKRLLQDLEAKVATSGSDPYIYNLNAMEGRAKEQVHNEVLYHVTDTMFQGGCWHAHR 720	
CC	DR;	InterPro; IPR001478;	PDZ.	Y	721 VNSYTMKDIAHGTTIPNYSRAQQOLIAQDMTQCVTRKSSGGPKQLVRYVSMDKAK 780	
CC	DR;	Pfam; PPF00595;	PDZ_1.	Db	721 VNSYTMKDIAHGTTIPNYSRAQQOLIAQDMTQCVTRKSSGGPKQLVRYVSMDKAK 780	
CC	DR;	SMART; SM00072;	GUKC_1.	Y	781 ASPLRLSPFDRGQDLPSPMGSSTICFWAESCLTVYPYLWPHPARPVLVPRAVGK 840	
CC	DR;	SMART; SM00238;	BDZ_1.	Db	781 ASPLRLSPFDRGQDLPSPMGSSTICFWAESCLTVYPYLWPHPARPVLVPRAVGK 840	
CC	DR;	PROSITE; PS50209;	CARD_1.	Y	781 ASPLRLSPFDRGQDLPSPMGSSTICFWAESCLTVYPYLWPHPARPVLVPRAVGK 840	
CC	DR;	PROSITE; PS00856;	GUANYLATE_KINASE_1;	Y	781 ASPLRLSPFDRGQDLPSPMGSSTICFWAESCLTVYPYLWPHPARPVLVPRAVGK 840	
CC	DR;	PROSITE; PS50052;	GUANYLATE_KINASE_2;	Y	781 ASPLRLSPFDRGQDLPSPMGSSTICFWAESCLTVYPYLWPHPARPVLVPRAVGK 840	
CC	DR;	PROSITE; PS50106;	PDZ_1.	Db	781 ASPLRLSPFDRGQDLPSPMGSSTICFWAESCLTVYPYLWPHPARPVLVPRAVGK 840	
CC	KW	colled coil.	CARD.	Y	841 LSEKLCILQGFKKCLAEYLSQEELYAWASORGDLIQEGETGGGRCWTRHAYESLMEMKNTH 900	
CC	FT	DOMAIN	15	107	FT	841 LSEKLCILQGFKKCLAEYLSQEELYAWASORGDLIQEGETGGGRCWTRHAYESLMEMKNTH 900
CC	FT	DOMAIN	128	409	FT	901 ALDDYQLDSYCTLHRMDIPIVIVHSVNEMKAKLKKKGQLGTSETEOLLEARQEEGDL 960
CC	FT	DOMAIN	568	658	FT	901 ALDDYQLDSYCTLHRMDIPIVIVHSVNEMKAKLKKKGQLGTSETEOLLEARQEEGDL 960
CC	FT	DOMAIN	858	990	FT	901 ALDDYQLDSYCTLHRMDIPIVIVHSVNEMKAKLKKKGQLGTSETEOLLEARQEEGDL 960
CC	FT	CONFLICT	619	671	FT	901 DRAPCLYSISLAPDGMSLDLQELSCVRQTADEQQKKVWTEQSPR 1004
CC	FT	AAH01326_1.			FT	901 DRAPCLYSISLAPDGMSLDLQELSCVRQTADEQQKKVWTEQSPR 1004
CC	FT	SEQUENCE	1004	AA:	FT	901 DRAPCLYSISLAPDGMSLDLQELSCVRQTADEQQKKVWTEQSPR 1004
CC	SQ	MW:	7468B8B56BE06073	CRC64_;	Db	961 DRAPCLYSISLAPDGMSLDLQELSCVRQTADEQQKKVWTEQSPR 1004

RESULT 2

CARE_MOUSE STANDARD PRT; 999 AA.

AC Q99KFO; DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Caspase recruitment domain protein 14 (Bcl10-interacting MAGUK Protein

DE 2) (Bimp2).

GN CARD14 OR BIMP2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TAXID=10909;

RN [1].

RN SEQUENCE FROM N.A.

RP MEDLINE:21391892; PubMed=11387339;

RC McAllister-Lucas L.M., Inohara N., Lucas P.C., Ruland J., Benito A.,

RA Li Q., Chen S., Chen F.F., Yamada S., Verma I.M., Mak T.W.,

RA Nunez G.;

RT "Bimp1, a MAGUK family member linking protein kinase C activation to

RT Bcl10-mediated NF-kappa B induction,"

J. Biol. Chem. 276:30597-30597(2001).

RL [2].

RN SEQUENCE OF 82-743 FROM N.A.

RP TISSUE=treat;

RC Strausberg R.;

RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBUNIT: CARD14 and Bcl10 bind to each other by CARD-CARD

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: CONTAINS 1 CARD DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 GUANYLATE KINASE-LIKE DOMAIN.

CC -1- CAUTION: Supposed to contain a SH3 domain which is not detected by

CC PROSITE, Pfam or SMART.

CC -----

----- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC entities

or send an email to license@lsb-sib.ch).	
CC	Db 716 HVNPNTMKDMEP-GTPNYSQQQOLLALIQDMQTCTVPRKPGPKLVRIVSVDK 773
CC	Db 779 AKASPIRLSFDRGOLDSRMEGS-STCFWAESCLTIVPTYLWPHPARPPVPLVPRAV 837
DR	Db 774 AAVSPITSSPDSQNSGKEGGPSVCEWSSESCFPLAPVLPVHPARPPVPLFPVRV 833
EMBL; EMBL; BC004692; AAHQ4692; 1;	Db 838 GKILSEKLCLLQGFKCLAEYLQSEYEANSQRGDIQEGEVSGGCRWTRHAYEMLMEK 897
InterPro; IPR001478; PDZ; 1;	Db 834 GRIGKIKLCLQGFKCSAEVLQSEQATWSQRGDDIQEGESIGDHWHWTRHAEELMMN 893
SMART; SMART; SM00228; PDZ; 1;	Db 898 NTHALLDVQLDSCVCLHMDIFPTVTHSYNEKMARKLKQQLRQTSCEFLLEARQEE 957
PROSITE; PS00209; CARD; 1;	Db 894 STHALDVRLDSVRHLRMDFPILTHSYNEKTAKLKRQSLHRSSEQFLEVARQEE 953
PROSITE; PS00836; GUANYLATE_KINASE; 1; FALSE_NEG.	Db 958 GDLDRAPCLYSSLAPGDWSLDGLLSCVROTADEEKVVWTEQSP 1003
PROSITE; PS50052; GUANYLATE_KINASE; 2; 1.	Db 954 GELDRVPCLYSSLAPGDWSLDLSLCVRAIADEKKVWTE-SP 998
DR	Db 999 AA: 113496 MW: D1B83500DA12410255 CRC64;
DR	Db 107 CARD. Coiled coil; DOMAIN 15 107 FT DOMAIN 125 411 FT DOMAIN 572 655 FT DOMAIN 854 986 FT CONFUNCT 736 743 SEQUENCE 999 AA: 113496 MW: D1B83500DA12410255 CRC64;
KW	Qy 1 MGEGLRRDSLTADEETLWEMMMSHRHRYVRICPSPRLTPYLRQAKVLQDDEEVLI 60 Db 1 MAELCRMDSLTADDEEMLWDMLSHRCRIVQSCPSRULPQAKVLQDDEEVLI 60 Qy 61 PRLTNSAMRAGHLLDLTKTGNGAIAFLSKLKHNPDPVTTVGLQDVFDSNFSLME 120 Db 61 SRFINSAMRAGHLLDLKARGKNGAIAFLSKLKHNPDPVTTVGLQSDIDFSFGSLME 120 Qy 121 TSKLTECLAGAIGSQQELNQEKGKEYVLRLRCCQOLQEHGLAETRAEGIHOLEADHSRM 180 Db 121 TSKLTECLAGAISSSLQEEQAKVEVLRLRCCQOLQERKIGLAAHAEGRQLEVDHSRM 180 Qy 181 KREYSAHFEVLRIKDEMELSLHYSNAQKEKLAASRSQELQEEQYLKQELORANMVS 240 Db 181 KREYSAHFEVLRIKDEMELSLHYSNAQKEKLAATRCHSLLQEEQYLKQELQRASLYS 240 Qy 241 SCELLEQOSELTSQESGDEELNRLKEENEKLRSLSLTAELQDILEQSDEARSROS 300 Db 241 SCERSRSRLLMASNLPEQEEELNRKENEKLSMSMTSLSVERDLSDEARESKS 300 Qy 301 LVERITHSLRERRVAALERQWEKEQTLQEQKSMACQOLYREKVNALQAQYCELOKE 360 Db 301 LVDRITHSLRERRVAALERQWEKEQTLQFRKTQDCELYKEKMTMQLQVAELQKE 360 Qy 361 RDQAYSARDQAQRETSQSLVEKSDSRQVEFELTDQCELETLQRLQAEPGPVULKQEAR 420 Db 361 RDQAYTARDQAQMEISQRLVKEWLKDALLRYYVEFQCLRTQLRQLQAPGGPKQEA 420 Qy 421 REPCPREKQVLRMQHAICPDDSDGLSYVSTESLSDLSATSSRELVFRSSPAPPS 480 Db 421 RELCIGKQVLVRMHAVCPDDSDLSSSTESLRWUDLSTSSEQMSFRSSSPMPMS 480 Qy 481 QOSLYKRVARDFGEPEPWSSFSCCLEPEGDGPALGAKAG--DPHDYELJLDTADLIPQLES 538 Db 481 QOSLYKRVARDFLEDPES---LSPEVLMRLOGATDQDLEFMIDGADSQTED 536 Qy 539 SLOQVSPGRDVSSEGVLMRPPARRILSQTVMIAFGQDALLEQIVGQNLTGFIRV 598 Db 537 SLOQ-SRSLSNVESSVPPRARKLTSQVTLAFQGQDALLEQIVGQNLTGFIRV 595 Qy 599 TPGSAQADMPLRGTOVIMMYDYEASEPLFLPEAVGLLRRVDFCCCLSVKYNT 658 Db 596 TPGSADEMALRPGTOVIMMYDKPTPKPSLATELNTLEAVGLLRRVNGSCYLSVKINT 655 Qy 659 DGYKRLQDLEAKVATSGDSFYIRVNLMGRKAGELYCHNEVLYHTDMEQGGCWH 718 Db 656 EGYNKLQDLDKAVYTSQDSEFYIRVNLMGRQGDDGELQTCNDLHVDTMFOGRSCWHA 715 Qy 719 HRVNSVYTMKIDAAHOTIPNVSRAOOQLLQDWTQCTVTRKBPSGGPQLRVTISMDK 778
Best Local Similarity 76.9%; Pred. No. 4e-186;	Score 3869 5; DB 1; Length 999;
Matches 774; Conservative 83; Mismatches 138; Indels 11; Gaps 7;	RESULT 3 CARB_HUMAN STANDARD; PRT; 1147 AA.
Qy 1 MGEGLRRDSLTADEETLWEMMMSHRHRYVRICPSPRLTPYLRQAKVLQDDEEVLI 60 Db 1 MAELCRMDSLTADDEEMLWDMLSHRCRIVQSCPSRULPQAKVLQDDEEVLI 60 Qy 61 PRLTNSAMRAGHLLDLTKTGNGAIAFLSKLKHNPDPVTTVGLQDVFDSNFSLME 120 Db 61 SRFINSAMRAGHLLDLKARGKNGAIAFLSKLKHNPDPVTTVGLQSDIDFSFGSLME 120 Qy 121 TSKLTECLAGAIGSQQELNQEKGKEYVLRLRCCQOLQEHGLAETRAEGIHOLEADHSRM 180 Db 121 TSKLTECLAGAISSSLQEEQAKVEVLRLRCCQOLQERKIGLAAHAEGRQLEVDHSRM 180 Qy 181 KREYSAHFEVLRIKDEMELSLHYSNAQKEKLAASRSQELQEEQYLKQELORANMVS 240 Db 181 KREYSAHFEVLRIKDEMELSLHYSNAQKEKLAATRCHSLLQEEQYLKQELQRASLYS 240 Qy 241 SCELLEQOSELTSQESGDEELNRLKEENEKLRSLSLTAELQDILEQSDEARSROS 300 Db 241 SCERSRSRLLMASNLPEQEEELNRKENEKLSMSMTSLSVERDLSDEARESKS 300 Qy 301 LVERITHSLRERRVAALERQWEKEQTLQEQKSMACQOLYREKVNALQAQYCELOKE 360 Db 301 LVDRITHSLRERRVAALERQWEKEQTLQFRKTQDCELYKEKMTMQLQVAELQKE 360 Qy 361 RDQAYSARDQAQRETSQSLVEKSDSRQVEFELTDQCELETLQRLQAEPGPVULKQEAR 420 Db 361 RDQAYTARDQAQMEISQRLVKEWLKDALLRYYVEFQCLRTQLRQLQAPGGPKQEA 420 Qy 421 REPCPREKQVLRMQHAICPDDSDGLSYVSTESLSDLSATSSRELVFRSSPAPPS 480 Db 421 RELCIGKQVLVRMHAVCPDDSDLSSSTESLRWUDLSTSSEQMSFRSSSPMPMS 480 Qy 481 QOSLYKRVARDFGEPEPWSSFSCCLEPEGDGPALGAKAG--DPHDYELJLDTADLIPQLES 538 Db 481 QOSLYKRVARDFLEDPES---LSPEVLMRLOGATDQDLEFMIDGADSQTED 536 Qy 539 SLOQVSPGRDVSSEGVLMRPPARRILSQTVMIAFGQDALLEQIVGQNLTGFIRV 598 Db 537 SLOQ-SRSLSNVESSVPPRARKLTSQVTLAFQGQDALLEQIVGQNLTGFIRV 595 Qy 599 TPGSAQADMPLRGTOVIMMYDYEASEPLFLPEAVGLLRRVDFCCCLSVKYNT 658 Db 596 TPGSADEMALRPGTOVIMMYDKPTPKPSLATELNTLEAVGLLRRVNGSCYLSVKINT 655 Qy 659 DGYKRLQDLEAKVATSGDSFYIRVNLMGRKAGELYCHNEVLYHTDMEQGGCWH 718 Db 656 EGYNKLQDLDKAVYTSQDSEFYIRVNLMGRQGDDGELQTCNDLHVDTMFOGRSCWHA 715 Qy 719 HRVNSVYTMKIDAAHOTIPNVSRAOOQLLQDWTQCTVTRKBPSGGPQLRVTISMDK 778	
Best Local Similarity 76.9%; Pred. No. 4e-186;	RESULT 3 CARB_HUMAN STANDARD; PRT; 1147 AA.
Matches 774; Conservative 83; Mismatches 138; Indels 11; Gaps 7;	CARB_HUMAN STANDARD; PRT; 1147 AA.
Qy 1 MGEGLRRDSLTADEETLWEMMMSHRHRYVRICPSPRLTPYLRQAKVLQDDEEVLI 60 Db 1 MAELCRMDSLTADDEEMLWDMLSHRCRIVQSCPSRULPQAKVLQDDEEVLI 60 Qy 61 PRLTNSAMRAGHLLDLTKTGNGAIAFLSKLKHNPDPVTTVGLQDVFDSNFSLME 120 Db 61 SRFINSAMRAGHLLDLKARGKNGAIAFLSKLKHNPDPVTTVGLQSDIDFSFGSLME 120 Qy 121 TSKLTECLAGAIGSQQELNQEKGKEYVLRLRCCQOLQEHGLAETRAEGIHOLEADHSRM 180 Db 121 TSKLTECLAGAISSSLQEEQAKVEVLRLRCCQOLQERKIGLAAHAEGRQLEVDHSRM 180 Qy 181 KREYSAHFEVLRIKDEMELSLHYSNAQKEKLAASRSQELQEEQYLKQELORANMVS 240 Db 181 KREYSAHFEVLRIKDEMELSLHYSNAQKEKLAATRCHSLLQEEQYLKQELQRASLYS 240 Qy 241 SCELLEQOSELTSQESGDEELNRLKEENEKLRSLSLTAELQDILEQSDEARSROS 300 Db 241 SCERSRSRLLMASNLPEQEEELNRKENEKLSMSMTSLSVERDLSDEARESKS 300 Qy 301 LVERITHSLRERRVAALERQWEKEQTLQEQKSMACQOLYREKVNALQAQYCELOKE 360 Db 301 LVDRITHSLRERRVAALERQWEKEQTLQFRKTQDCELYKEKMTMQLQVAELQKE 360 Qy 361 RDQAYSARDQAQRETSQSLVEKSDSRQVEFELTDQCELETLQRLQAEPGPVULKQEAR 420 Db 361 RDQAYTARDQAQMEISQRLVKEWLKDALLRYYVEFQCLRTQLRQLQAPGGPKQEA 420 Qy 421 REPCPREKQVLRMQHAICPDDSDGLSYVSTESLSDLSATSSRELVFRSSPAPPS 480 Db 421 RELCIGKQVLVRMHAVCPDDSDLSSSTESLRWUDLSTSSEQMSFRSSSPMPMS 480 Qy 481 QOSLYKRVARDFGEPEPWSSFSCCLEPEGDGPALGAKAG--DPHDYELJLDTADLIPQLES 538 Db 481 QOSLYKRVARDFLEDPES---LSPEVLMRLOGATDQDLEFMIDGADSQTED 536 Qy 539 SLOQVSPGRDVSSEGVLMRPPARRILSQTVMIAFGQDALLEQIVGQNLTGFIRV 598 Db 537 SLOQ-SRSLSNVESSVPPRARKLTSQVTLAFQGQDALLEQIVGQNLTGFIRV 595 Qy 599 TPGSAQADMPLRGTOVIMMYDYEASEPLFLPEAVGLLRRVDFCCCLSVKYNT 658 Db 596 TPGSADEMALRPGTOVIMMYDKPTPKPSLATELNTLEAVGLLRRVNGSCYLSVKINT 655 Qy 659 DGYKRLQDLEAKVATSGDSFYIRVNLMGRKAGELYCHNEVLYHTDMEQGGCWH 718 Db 656 EGYNKLQDLDKAVYTSQDSEFYIRVNLMGRQGDDGELQTCNDLHVDTMFOGRSCWHA 715 Qy 719 HRVNSVYTMKIDAAHOTIPNVSRAOOQLLQDWTQCTVTRKBPSGGPQLRVTISMDK 778	
SEQUENCE FROM N.A.	SEQUENCE FROM N.A.
RX MEDLINE=21192234; PubMed=11278892;	RX MEDLINE=21192234; PubMed=11278892;
RA Bertin J., Wang L., Guo Y., Jacobson M.D., Poyet J.-L., Srinivasula S.M., Meirion S., Di Stefano P.S., Alnemri E.S.;	RA Bertin J., Wang L., Guo Y., Jacobson M.D., Poyet J.-L., Srinivasula S.M., Meirion S., Di Stefano P.S., Alnemri E.S.;
RT CARD11 and CARD14 are novel caspase recruitment domain members	RT CARD11 and CARD14 are novel caspase recruitment domain members
DE Caspase recruitment domain in protein 11 (CARD-containing MAGUK protein	DE Caspase recruitment domain in protein 11 (CARD-containing MAGUK protein
DE (Carma 1).	DE (Carma 1).
GN CARMA1.	GN CARMA1.
OS Homo sapiens (Human)	OS Homo sapiens (Human)
OC Metazoa; Chordata; Craniata; Euteleostomi;	OC Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.	OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TAXID=9606;	NCBI_TAXID=9606;
RN [1]	RN [1]
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RX PubMed=1136195;	RX PubMed=1136195;
RA Martiron F., Micheau O., Bonnet D., Thome M., Tschochopp J.;	RA Martiron F., Micheau O., Bonnet D., Thome M., Tschochopp J.;
RT Carmel, a CARD-containing binding partner of Bcl110, induces Bcl110 phosphorylation and NF-kappaB activation";	RT Carmel, a CARD-containing binding partner of Bcl110, induces Bcl110 phosphorylation and NF-kappaB activation";
RT Phosphorylation and NF-kappaB activation";	RT Phosphorylation and NF-kappaB activation";
RL J. Biol. Chem. 276:11877-11882(2001).	RL J. Biol. Chem. 276:11877-11882(2001).
RP ERRATUM.	RP ERRATUM.
RA Gaide O., Martiron F., Micheau O., Bonnet D., Thome M., Tschochopp J.;	RA Gaide O., Martiron F., Micheau O., Bonnet D., Thome M., Tschochopp J.;
RL PubMed=1136195;	RL PubMed=1136195;
CC -!!- FUNCTION: Activates NF-kappaB via Bcl110 and IKK. Stimulates the phosphorylation of Bcl110.	CC -!!- FUNCTION: Activates NF-kappaB via Bcl110 and IKK. Stimulates the phosphorylation of Bcl110.
CC -!!- SUBUNIT: CARD11 and Bcl110 bind to each other by CARD-CARD interaction.	CC -!!- SUBUNIT: CARD11 and Bcl110 bind to each other by CARD-CARD interaction.
CC -!!- SUBCELLULAR LOCATION: Cytoplasmic.	CC -!!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!!- TISSUE SPECIFICITY: Detected in adult peripheral blood leukocytes, thymus, spleen and liver. Also found in promyelocytic leukemia HL-60 cells, chronic myelogenous leukemia K562 cells, Burkitt's lymphoma Raji cells and colorectal adenocarcinoma SW480 cells. Not detected in Hela S3, Molt-4, A549 and G431 cells.	CC -!!- TISSUE SPECIFICITY: Detected in adult peripheral blood leukocytes, thymus, spleen and liver. Also found in promyelocytic leukemia HL-60 cells, chronic myelogenous leukemia K562 cells, Burkitt's lymphoma Raji cells and colorectal adenocarcinoma SW480 cells. Not detected in Hela S3, Molt-4, A549 and G431 cells.
CC -!!- SIMILARITY: CONTAINS 1 CARD DOMAIN.	CC -!!- SIMILARITY: CONTAINS 1 CARD DOMAIN.
CC -!!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.	CC -!!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -!!- CAUTION: Supposed to contain a SH3 domain which is not detected by PROSITE, Pfam or SMART.	CC -!!- CAUTION: Supposed to contain a SH3 domain which is not detected by PROSITE, Pfam or SMART.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstations on the European Bioinformatics Institute. There are no restrictions on	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstations on the European Bioinformatics Institute. There are no restrictions on

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC DR EMBL: AF322641; AAG3402; 1; -
CC DR InterPro: IPR01478; PDZ.
CC PFam: PF00595; PDZ; 1.
CC DR SMART: SM00228; PDZ; 1.
CC DR PROSITE: PS50209; CARD; 1;
CC PROSITE: PS50056; GUANYLATE_KINASE_1; FALSE_NEG.
CC DR PROSITE: PS50052; GUANYLATE_KINASE_2; FALSE_NEG.
CC DR PROSITE: PS50106; PDZ; FALSE_NEG.
CC DR Coiled coil.; CARD.
CC FT DOMAIN 11 103 COILED COIL (POTENTIAL).
CC FT DOMAIN 123 442 PDZ.
CC FT DOMAIN 673 748 GUANYLATE KINASE.
CC FT DOMAIN 966 1133 GUANYLATE KINASE.
CC FT CONFLICT 808 808 P -> L (IN REF. 2).
CC FT SEQUENCE 1147 AA: 132641 MW: 9134B015D2B36CC CRC64;
CC SQ
Query Match 24.18; Score 1239.5; DB 1; Length 1147;
Best Local Similarity 30.5%; Pred. No. 9.5e-55;
Matches 354; Conservative 204; Mismatches 402; Indels 201; Gaps 33;
Qy 15 DEETLWEMMESHRRHTRVRCIPSRLTPYLQRQKVLQCLDEEYLHSPLRTNSAMRAGHL 74
Db 11 EEDALWENVECNRHMLSRVINFPAKLTPTYLOCKVIDEDEVLNAPLPSKINRAGHL 70
Qy 75 DILKTRKGNAKAFLESIKFHNPDVTLTGTQPDYDSNFSGLMETSKLTCELAGTAGS 134
Db 71 DILHTKGQRGYYVFVFLSIEFYPPELYKLVTKGEPTTRFSTIVVEGHEGLTHELMNEVTK 130
Qy 135 LQEELNQEKGOKEVLLRCCQOLO-EHILGAAETRAEGLHQLEADHSRMKREVAHFHEVLR 193
Db 131 LOOMKMKDQLQELLARKRQLEDEKKRQEVTE-LITFOQRYKMKKEERDYSNDKLW 189
Qy 194 LKDEMILSLIHSNALQEKEKLAASRCRSIQQEELYLQKELQRANMVSSCQELELOEQSLSR 253
Db 190 VKDNDNYNIAMLRQAOLSEKINMAYMRSDRQLEIDQKLHRLNK--MEEEKLE-RNOSLKL 246
Qy 254 ASDOES--GDEELNRLKBNENKEPLSISI-----AEKDILEQSLDFAR 295
Db 247 KNDIENPRKEQVLELERENEMLKTKNOELQSIQIAQKRSPLPDSKATDILDELDKREAL 306
Qy 296 GSQELVETHHSRERAVAERAOREQWEKEQTLLQFSKSMACQLYREKVNALQAQVC 355
Db 307 EDQELDQYASRDSAQRBISQSLIVEKDELROVFEIQLKCEKDLLEKCTSLKGOCIMEYHRMTMVLQE 366
Qy 356 ELQERDQYASRDSAQRBISQSLIVEKDELROVFEIQLKCEKDLLEKCTSLKGOCIMEYHRMTMVLQE 401
Db 367 EVERERDQFHRSRDEAQTYQSQCILIEKDVKYRKQIRELEKNDRIEMYRREACTIVNLIES 426
Qy 402 QLRQKAE-----PPGYLKO-----EARTRPCPREK----- 4.28
Db 427 KLRRISKDSNNLDOSLPRNLPVTIISODGQDASPRNTNGQADDSTSSESPDSSKYFLPY 4.86
Qy 429 ---ORLYRMHAI-CPRDDSDCSLVSSTEQOLL-----SPDLSATSS-SRRLVDSPRSSSP 476
Db 487 HPPORRMKNGIQLQRKAPBISPLTQKAGHDEEGTDAASPSSCGSLPTNSSTKMQP 546
Qy 477 -----APPSQSOQSLYKRYAEDFGEPWSFSSCLEIPEGDPGALPGAKA-GDPHL 523
Db 547 PRSRSSIMMSTAEPFGNDDSVTRRYKEDAPHR-----STVE-EDNSGGFDALDDDSHE 600
Qy 524 DY-----ELLDTADLPQLESSLQPVs---EARTRPCPREK-----
Db 601 RYSGCPSSSIHSSSHQSESSDAYLEONLMLFRKFSLPERPSPVTSVHVRGPSPS-- 658
Qy 565 ILSQLQTMIAFGDALLEQISVQGNNITGIFTTHRGPSAADQMAJRPQTIVMDVEASE 624
Db 659 ----VQHTTLNGDSLTSQTLIGGNGARSFVHSVKPSLAEKAGLREGHQLLLEGCIRG 714

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	SEQUENCE FROM N.A.
RP	SPAIN=129;
RC	MEDLINE=93252986; PubMed=8486731;
RX	Itch M., Nagafuchi A., Yonemura S., Yasuda-Kitani T., Tsukita S., Tsukita S.;
RA	The 220-kD protein colocalizing with cadherins in non-epithelial cells is identical to ZO-1, a tight junction-associated protein in epithelial cells: cDNA cloning and immunoelectron microscopy.";
RT	J. Cell Biol. 121:491-502(1993).
RT	-1- FUNCTION. THE N-TERMINAL MAY BE INVOLVED IN TRANSDUCING A SIGNAL REQUIRED FOR TIGHT JUNCTION ASSEMBLY. WHILE THE C-TERMINAL MAY HAVE SPECIFIC PROPERTIES OF TIGHT JUNCTIONS. THE ALPHA DOMAIN MIGHT BE 3).
RT	-1- SUBUNIT: INTERACTS WITH OCCLUDIN, CLAUDINS AND ZO-3.
RT	-1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE. CYTOPLASMIC SIDE. MOVEMENT OF ZO-1 FROM THE CITOPLASM TO MEMBRANE IS AN EARLY EVENT OCCURRING CONCURRENTLY WITH THE CYTOSKELETON-CELL CONTACT (BY SIMILARITY).
RL	-1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC	-1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC	-1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; D1340; BAA03274_1; DR PIR; AA6431; A4631.
CC	DR HSSP; P31016; 1BEG.
CC	DR MGII; 98/59; Tjpl.
CC	DR InterPro; IPR00619; Guanylate_kin.
CC	DR InterPro; IPR001478; PDZ.
CC	DR InterPro; IPR001452; SH3.
CC	DR InterPro; IPR000506; ZU5.
CC	DR Pfam; PF00625; Guanylate_kin; 1.
CC	DR Pfam; PF00055; PDZ; 3.
CC	DR Pfam; PF00018; SH3; 1.
CC	DR Pfam; PF00791; ZU5; 1.
CC	DR SMART; SM00072; GuKc; 1..
CC	DR SMART; SM00248; PDZ; 3.
CC	DR SMART; SM00326; SH3; 1.
CC	DR SMART; SM00218; ZU5; 1.
CC	DR PROSITE; PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
CC	DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
CC	DR PROSITE; PS50106; PDZ; 3.
CC	DR PROSITE; PS50002; SH3; 1.
KW	Tight junction; SH3 domain; Repeat; Membrane.
DB	403 EAQ--LLAEGIQLRR--QQETLVSSLDFEDGSPRSQEELDQLSKGCLA 456
DB	FT DOMAIN 23 110 PDZ 1.
DB	FT DOMAIN 186 264 PDZ 2.
DB	FT DOMAIN 421 502 PDZ 3.
DB	FT DOMAIN 516 584 SH3.
DB	FT DOMAIN 644 794 GUANYLATE KINASE.
DB	FT DOMAIN 1242 1247 POLY-PRO.
DB	FT DOMAIN 1424 1430 POLY-PRO.
SQ	SEQUENCE 1745 AA; 194710 MW; C3DA2C0A9F411F66 CRC64;
Query	Match 5.6%; Score 289; DB 1; Length 1745;
Query	Best Local Similarity 22.3%; Pred. No. 4.3e-07;
Matches	191; Conservative 136; Mismatches 319; Indels 212; Gaps 43;
QY	231 QELQRANNVSSCELELOFQSLSRATDQSDEFLNRLKEENELRSLTFSLAEKDILEQS 290
Db	70 QENDRVAVNGYSMDNVHAFQQLRKSGKNA -KITIRRKKKVQIPVSHDPPEVSDN 127
QY	291 LDEARGSSQELVERIHSRE -RAVAAPRQEQQWEKEQTLLQQFKSKMACQLYREKVN 349
Db	128 EDDSYD-----BEVHDPRAGRGALANRSEWARDRSA --SRERSLSPRSDRRSVAS 178

light as well as adherens junctions with a binding affinity to

RT tight as well as adherens junctions with a binding affinity to

RT occludin and alpha catenin;

RL J. Biol. Chem. 274:5981-5986(1999);

CC -!- FUNCTION: PLAYS A ROLE IN TIGHT JUNCTIONS AND ADHERENS JUNCTIONS.

CC -!- SUBUNIT: INTERACTS WITH OCCLUDIN.

CC -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE. CYTOPLASMIC SIDE.

CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.

CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF113005; ADD1996_1; -

HSSP; P31016; 1BE9.

MGD; MGI:1341872; TJP2.

DR InterPro; IPR006169; Guanylate_kin.

DR InterPro; IPR01418; PDZ.

DR InterPro; IPR001452; SH3.

DR Pfam; PF00925; Guanylate_kin; 1.

DR Pfam; PF00595; PDZ; 3.

DR Pfam; PF00018; SH3; 1.

DR SMART; SM00072; GuKc; 1.

DR SMART; SM00228; PDZ; 3.

DR SMART; SM00326; SH3; 1.

DR PROSITE; PS00856; GUANYLATE_KINASE_1; FALSE_NEG.

DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.

DR PROSITE; PS50106; PDZ; 3.

DR PROSITE; PS50002; SH3; 1.

KW Tight junction; SH3 domain; Repeat; Membrane.

FT DOMAIN 10 97 PDZ 1.

FT DOMAIN 287 365 PDZ 3.

FT DOMAIN 489 570 PDZ 3.

FT DOMAIN 584 649 SH3.

FT DOMAIN 678 858 GUANYLATE KINASE.

FT DOMAIN 1139 1142 POLY GLU.

SQ SEQUENCE 1167 AA; F15DA3EB3F9434F CRC64;

Query Match 5.4%; Score 280; DB 1; Length 1167;

Best Local Similarity 23.0%; Pred. No. 7-70-07;

Matches 199; Conservative 122; Mismatches 308; Indels 234; Gaps 43;

QY 700 NEVLHVTDMFQG-CGCMWAHVNSYTMDTAAGHTIPNYSRAQQQLIALIDMTQCTV 758

DB 540 GEFRVYDTLYNGKLGSWLAIRGKHNKE--VERGLIPKNKRAEQ--LASTVQ-----Y 588

QY 759 TRKPGGGPKLVRYVSMNDKAKASPLR-LSFDRGQDPSRMEGSSGTCAWEASCLTLYP-Y 816

DB 589 TLPKTAGG--DRADEFWRFLGRSSKRNLRSKRSRDL--AQPVQTKEPAY 633

QY 817 TLWPHRPARPRLVPLVPRAVKGKILSEKLQQFKKCLAEYLSQLEE--YE-AWSORGD 872

DB 634 ERVVLREAGFLRPTIF-GPIADVAIRE--LAREEDIYQIAKSEIRD 678

QY 873 IIOGEYSGGRCRMVTHAVESLMEKNTHALDVFOLDSVCTLHMRDIFPIVIVHSNEKMA 932

DB 679 AGTDHB RSSG--IIRLHTIKQIQIDQDKHALDVTPNAYDRNAYAQWPIVVLNPDSKQG 735

QY 933 KKL-----KKGKLRGLTSBQEELAARQEGDLDRAPCLYSSAPDGFSLDGLIS 983

DB 736 VKTMRMRLCPESRKSARKLYERSHKL----RKNNHHLFTTITLNMSM-NDGW-----YG 784

QY 984 CYRQAIADBKQVWTEQ 1001

DB 785 ALKEAIQQQNQLVWVE 802

RESULT 9

ZO2_MOUSE STANDARD; PRNT; 1167 AA.

ID ZO2_MOUSE STANDARD; PRNT; 1167 AA.

AC Q92AU1; ID ZO2_MOUSE STANDARD; PRNT; 1167 AA.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Tight junction protein ZO-2 (Zonula occludens 2 protein) (Zona occludens 2 protein) (Tight junction protein 2).

GN TJP2 OR ZO2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBItaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6 X CBA;

RD MEDLINE=99150392; PubMed=10026224;

RX Itoh M., Morita K., Tsukita S.; Characterization of ZO-2 as a MAGUK family member associated with

|| - SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.

CC CC - This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL: L27152; AAC37332_1; -.

CC DR HSSP: P31016; IBB3.

CC DR InterPro: IPR0006619; Guanylate_Kin.

CC DR InterPro: IPR01478; PDZ.

CC DR InterPro: IPR001452; SH3.

CC DR Pfam: PF00625; Guanylate_kin; 1.

CC DR Pfam: PF00595; PDZ; 3.

CC DR SMART: SM00072; GuKc; 1.

CC DR SMART: SM00028; PDZ; 3.

CC DR SMART: SM0026; SH3; 1.

CC DR PROSITE: PS00856; GUANYLATE_KINASE_1; FALSE_NEG.

CC DR PROSITE: PS50052; CUANYLATE_KINASE_2; 1.

CC DR PROSITE: PS50106; PDZ; 3.

CC DR PROSITE: PS50002; SH3; 1.

CC KW Tight junction, SH3 domain; Repeat; Membrane; Alternative splicing.

CC ET DOMAIN 10 97 PDZ 1.

CC ET DOMAIN 291 369 PDZ 2.

CC FT DOMAIN 493 574 PDZ 3.

CC FT DOMAIN 588 653 SH3.

CC FT DOMAIN 680 860 GUANYLATE KINASE.

CC SQ SEQUENCE 1174 AA; 132085 MW; 45ABB36BDBB1226 CRC64;

Query Match 5.0%; Score 256; DB 1; Length 1174;

Best Local Similarity 22.9%; Pred. No. 1.2e-05;

Matches 196; Conservative 124; Mismatches 325; Indels 210; Gaps 46;

QY 210 QEEELAASCRCSRQEEEL---YLLKQPLQRANMVSSCCELEQESQGEEL 264

Db 164 QEEE--RSGRSLERGLHDHDDDRRPRERSRG-----RSERGLDH--DDY 206

QY 265 NRLKEENEKLR-----LTFSLAEKDILEQSLEAR--GSRQELLVERI 305

Db 207 GREGERSHGMSTURGYDGYDRGDYRTYS-PFEEYGRFTQDPRHAGSRSR--SRE 263

QY 306 HSURERAAE-QREQWEEKTLLQFQKSXMACOLYREKVNALAQCVCELOKERDQA 364

Db 264 H-LRSRSPSPSLRGRPHAGQPD-----SDRPIGVLLMKSAN-----EE 302

QY 365 YSDASDAREISO---SLVEKDSLRQ-----LRTQIROLQAEPPVILK 415

Db 303 YGURLGSQFIKOMTRPLATDGNLHEGDIILKINGTVTENMSLTARKLIEKSRLK- 361

QY 416 QEARTREPCPREKLYRMHAICPRDDSDCSL--VSSTES-----QLISDLSAT5 463

Db 362 QLVLRD----SKQTLNIPSL---NDSDESEDSIEENRSFSPERRQQSYDYS 414

QY 464 SRLVLDSEFSSPAPPSQOSLYKRAEDFGEEPWFSSCCLEPEGDQALPGAK-AGDPH 522

Db 415 SNEKL-----KERPNNSEDMQNWRSR-MGATPTPFSM-----GDIASVYCTENSEKPR 462

QY 523 LDYELDITADLPOLESSQSPGSRPARRFL--SCVTMLAF-QGDAL 579

Db 463 YQEE----PAQPKAAPP-----TFLRSPSPEDAYGPNTKMYRFKKGDSV 505

QY 580 LEQISVIGGNLTQPSGSAQDMALRPGTQIYMDYEASEPLEFKAVLEDTTLREA 639

RT RT PARTIAL SEQUENCE FROM N.A.
MEDLINE=94179414; PubMed=8132716;

RA RA Beatch M., Jesaitis L.A., Goodenough D.A., Stevenson B.R.;
"The tight junction protein ZO-2 contains three PDZ (PSD-95/Discs-ZO-1) domains and an alternatively spliced region.";
J. Biol. Chem. 271:25723-25726(1996).

RL RN [2] PARTIAL SEQUENCE FROM N.A.
MEDLINE=94179414; PubMed=8132716;

RT RT PARTIAL SEQUENCE FROM N.A.
"Molecular characterization and tissue distribution of ZO-2, a tight junction protein homologous to ZO-1 and the Drosophila discs-large tumor suppressor protein";
J. Cell Biol. 124:949-961(1994).

CC CC -1- FUNCTION: PLAYS A ROLE IN TIGHT JUNCTIONS AND ADHERENS JUNCTIONS.
"SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE, CYTOPLASMIC SIDE."
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPlicing.

CC CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
-1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.

Db 559 - -LLEIPKGEMVTILAOSRADDVYRDIL-----ACGRGDSEFFRSHECEKTPQSLSAFS 610
QY 699 CNEVLHVDTMFMQG-CGCMVHARVNSYTMKDTIAHGTLPNYSRAQQQLITALIDDMTQQCT 757
Db 611 RGEYERVVDTLYDGLKLGHWLAKRIGNELEK----GLIPNKSRAEQ----MASTQN----- 657
QY 758 VTRPKSSGGPQKLVRIVSMKDCKASAPLSDFRQLDPSRMGSSTCFWAESCLTLVPT 817
Db 658 -AQDRNAGDRADFWRM---RGQRSGMKKKNLNRISREDLTAAVSVTKPPA-----YE 704
QY 818 LYMPHRPARPRPVILL---VPRAVGKVLSEKULLQGFKKCLAEYLQQEEYAWSQRGDI 873
Db 705 RVLLREAGFKRPVLFGPPIADIALEKLANELPDLFQPTAK-----TEPKDAGEKSSG 756
QY 874 IQEDEVSGGRCWTRHAYESLMEKNTHALDVQLDSVCTLHHMDIFPVIVHVSVERNAK 933
Db 757 V-----VRLNTYRQIICDQKHALLDVTPKAIDLNNYTQWPVIVFNPDSDROGV 805
QY 934 KLLKGLQRUTSEQQLEEARQBERGDLDRAPLY-----SSLAPDGWSLDLGLLSCVR 986
Db 806 KTHR--QRLNPTSNSKSSRKLYDOANKLKT-CAHLFTATINLSANDSW-----FGSLK 856
QY 987 QAIADEQKKKVWTEQ 1001
Db 857 DTIQHQGEAWWSE 871

RESULT 12

ZO2_HUMAN STANDARD: PRT; 1190 AA.
 ID ZO2_HUMAN Q9UDY2; Q9UDY1; Q9UDY0; Q99839;
 AC Q9UDY2; Q9UDY1; Q9UDY0; Q99839;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Right junction protein 20-2 (zonula occludens 2 protein) (Zona occludens 2 protein) (Tight junction protein 2).
 GN TJP2 OR ZO2 OR X104.
 OS Homo sapiens (Human).
 OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrini; Hominidae; Homo.
 RN NCBI_TAXID=9606;
 RN OX

SEQUENCE FROM N.A. (ISOFORM A1).

RC TISSUE=brain;
 RC MEDLINE=95038744; PubMed=7551235;
 RA Duciatis F., Rodius F., Wrogemann K., Mandel J.-L., Koenig M.;
 RA "The Friedreich ataxia region: characterization of two novel genes and reduction of the critical region to 300 kb.";
 RL Hum. Mol. Genet. 3:909-914(1994).
 [2]

SEQUENCE FROM N.A. (ISOFORMS A1; C1; A2 AND C2).

RC TISSUE=Pancreas;
 RX MEDLINE=20472048; PubMed=11018256;
 RA Chlenski A., Ketels K.V., Korovaitseva G.I., Talamonti M.S., Oyasu R., Scarpellini D.G.;
 RA "Organization and expression of the human zo-2 gene (tjp-2) in normal and neoplastic tissues.";
 RL Biochim. Biophys. Acta 1493:319-324(2000).
 [3]

SEQUENCE OF 1-104 FROM N.A. (ISOFORMS A1 AND C1).

RC TISSUE=Pancreas;
 RX MEDLINE=99207578; PubMed=10360833;
 RA Chlenski A., Ketels K.V., Tsao M.-S., Talamonti M.S., Anderson M.R., Oyasu R., Scarpellini D.G.;
 RA "Tight junction protein zo-2 is differentially expressed in normal pancreatic ducts compared to human pancreatic adenocarcinoma.";
 RL Int. J. Cancer 80:137-144 (1999).
 [4]

PARTIAL SEQUENCE FROM N.A. (ISOFORMS A1 AND C1).

RC TISSUE=Pancreas;
 RX MEDLINE=99446875; PubMed=10495477;
 RA Chlenski A., Ketels K.V., Engeriziser J.L., Talamonti M.S., Tsao M.-S., Koutnikova H., Oyasu R., Scarpellini D.G.;

RT "zo-2 gene alternative promoters in normal and neoplastic human pancreatic duct cells";
 RL Int. J. Cancer 83:349-358(1999).
 RN [5]
 RP SEQUENCE OF 1047-1167 FROM N.A.
 RC TISSUE=Aortic smooth muscle;
 RA Adams L.D., Werny I., Schwartz S.M.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: PLAYS A ROLE IN TIGHT JUNCTIONS AND ADHERENS JUNCTIONS.
 CC -!- SUBUNIT: INTERACTS WITH OCCLUDIN.
 CC -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE. CYTOPLASMIC SIDE (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A1 (SHOWN HERE), C1, A2 AND C2;
 CC -!- ARE PRODUCED BY ALTERNATIVE SPlicing AND ALTERNATIVE INITIATION.
 CC -!- TISSUE SPECIFICITY: THIS PROTEIN IS FOUND IN EPITHELIAL CELL.
 CC JUNCTIONS. ISOFORM A1 IS ABUNDANT IN THE HEART AND BRAIN WHEREAS ISOFORM C1 IS EXPRESSED AT HIGH LEVEL IN THE KIDNEY, PANCREAS, HEART AND PLACENTA. IN BRAIN AND SKELETAL MUSCLE ONLY ISOFORM A1 IS DETECTABLE. ISOFORM C1 IS FOUND IN NORMAL AS WELL AS IN MOST NEOPLASTIC TISSUES WHILE ISOFORM A1 IS PRESENT ALMOST EXCLUSIVELY IN NORMAL TISSUE.
 CC -!- SIMILARITY: CONTAINS 3 PDZ/DFR DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
 CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITION 1086 ONWARD DUE TO A FRAMESHIFT.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR EMBL; L27476; AAA61300_1; ALT_FRAME.
 CC DR EMBL; AF043195; ADD20387_2;
 CC DR EMBL; AF043196; ADD20387_2; JOINED.
 CC DR EMBL; AF043197; ADD20387_2; JOINED.
 CC DR EMBL; AF177518; ADD20387_2;
 CC DR EMBL; AF177519; ADD20387_2; JOINED.
 CC DR EMBL; AF177520; ADD20387_2; JOINED.
 CC DR EMBL; AF177521; ADD20387_2; JOINED.
 CC DR EMBL; AF177522; ADD20387_2; JOINED.
 CC DR EMBL; AF177523; ADD20387_2; JOINED.
 CC DR EMBL; AF177524; ADD20387_2; JOINED.
 CC DR EMBL; AF177525; ADD20387_2; JOINED.
 CC DR EMBL; AF177526; ADD20387_2; JOINED.
 CC DR EMBL; AF177527; ADD20387_2; JOINED.
 CC DR EMBL; AF177528; ADD20387_2; JOINED.
 CC DR EMBL; AF177529; ADD20387_2; JOINED.
 CC DR EMBL; AF177530; ADD20387_2; JOINED.
 CC DR EMBL; AF177531; ADD20387_2; JOINED.
 CC DR EMBL; AF177532; ADD20387_2; JOINED.
 CC DR EMBL; AF177533; AAC02527_2;
 CC DR EMBL; AF043196; AAC02527_2; JOINED.
 CC DR EMBL; AF043197; AAC02527_2; JOINED.
 CC DR EMBL; AF177518; AAC02527_2; JOINED.
 CC DR EMBL; AF177519; AAC02527_2; JOINED.
 CC DR EMBL; AF177520; AAC02527_2; JOINED.
 CC DR EMBL; AF177521; AAC02527_2; JOINED.
 CC DR EMBL; AF177522; AAC02527_2; JOINED.
 CC DR EMBL; AF177523; AAC02527_2; JOINED.
 CC DR EMBL; AF177524; AAC02527_2; JOINED.
 CC DR EMBL; AF177525; AAC02527_2; JOINED.
 CC DR EMBL; AF177526; AAC02527_2; JOINED.
 CC DR EMBL; AF177527; AAC02527_2; JOINED.

Db	811	REFNPDSROGVKTMR-QRLNPTSNKSSRKLEDDANKL-----KKTCAHLFATINLNSA	863	INTERMEDIATE FILAMENTS DURING MITOSIS (BY SIMILARITY); DISEASE: DEFECTS IN PLECTIN ARE THE CAUSE OF EPIDERMOLYSIS BULLOSA WITH MUSCULAR DYSTROPHY (MD-EBS OR EB-MD), AN AUTOSOMAL RECESSIVE DISORDER CHARACTERIZED BY EPIDERMAL BLISTER FORMATION AT THE LEVEL OF THE HEMIDESMOSOME AND ASSOCIATED LATE-ONSET MUSCULAR DYSTROPHY.
QY	972	PDRGSDLDGLLSCVROTADEQQKVVTEQ	1001	-1- SIMILARITY: CONTAINS 1 ACTIN BINDING DOMAIN.
Db	864	NDSW-----FGSLKDQTQHQGEAWVSE	887	-1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS. -1- SIMILARITY: CONTAINS 3 PLECTIN REPEATS.
RESULT	13	PLEC_HUMAN	STANDARD;	-1- SIMILARITY: CONTAINS 4 PLECTIN REPEATS. BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.
ID		PLEC_HUMAN	PRT;	4684 AA.
RN		Q15149; Q16640; Q15148;		CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
AC		DT 16-OCT-2001 (Rel. 40, Last sequence update)		CC -1- CC
RC		DT 16-OCT-2001 (Rel. 40, Last annotation update)		CC -1- CC
RA		DE PLEC1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1).		CC -1- CC
GN		Homo sapiens (Human).		CC -1- CC
OS		OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;		CC -1- CC
OC		OC Mauroyota; Metazoa; Primates; Homidae; Homo.		CC -1- CC
OX		OX NCBI_TaxID=9606;		CC -1- CC
RN		[1] SEQUENCE FROM N.A. (ISOFORM 1).		CC -1- DR EMBL; U53204; AAB05427; 1; -.
RC		RC TISSUE-Placenta;		CC -1- DR EMBL; U63610; AAB05428; 1; -.
RA		RA MEDLINE-#9621066; PubMed-8633055;		CC -1- DR EMBL; U63609; AAB05429; 1; -.
RA		RA Liu C.-G., Maercker C., Castanon M.J., Hauptmann R., Wiche G.; RT "Human plectin: organization of the gene, sequence analysis, and chromosome localization (8q44)." Proc. Natl. Acad. Sci. U.S.A. 93:4278-4283(1996).		CC -1- DR EMBL; X97053; CAA55765; 1; -.
RA		RA RT [2] SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND DISEASE.		CC -1- DR HSSP; Q01082; 1BKR.
RA		RA MEDLINE-#96312447; PubMed-8638233;		CC -1- DR MM; 601282; -.
RA		RA Bullrich F., Pulkkinen L., Smith F.J.D., Rugg E.L., Lane E.B., McLean W.H.I., Burgesson R.E., Amaro S., Hudson D.L., Owaribe K., McGrath J.A., McMillan J.R., Eady R.A.J., Leigh I.M., Christiano A.M., Utton J.; RT "Loss of plectin causes epidermolysis bullosa with muscular dystrophy." CDNA cloning and genomic organization."; RT Genes Dev. 10:1724-1735(1996).		CC -1- DR InterPro; IPR001589; Actinin_act_bind.
RA		RA RT [3] VARIANT MD-EBS 1003-GLN-ALA-1005 DEL.		CC -1- DR InterPro; IPR001715; Calponin_hom.
RA		RA MEDLINE-#97049959; PubMed-8394687;		CC -1- DR InterPro; IPR001101; Plectin_repeat.
RA		RA Hachisuka H., Nishikawa T., McLean W.H.I., Utton J.; RT "Homozygous deletion mutations in the plectin gene (PLEC1) in patients with epidermolysis bullosa simplex associated with late-onset muscular dystrophy"; RT Gen. Mol. Genet. 5:1539-1546(1996).		CC -1- DR Pfam; PF00307; CH; 2.
RA		RA RT [4] VARIANT MD-EBS LEU-429 INS.		CC -1- DR SMART; SM00020; PLEC; 3.3.
RA		RA MEDLINE-#21090821; PubMed-11159198;		CC -1- DR SMART; SM00150; SPEC; 5.
RA		RA Bauer J.W., Rouan F., Kofler B., Reznicek G.A., Kornacker L., Wiche G., Hantner R., Klausenberger A., Huber A., Pohla-Gubo G., RT "A compound heterozygous one amino-acid insertion/nonsense mutation in the plectin gene causes epidermolysis bullosa simplex with plectin deficiency"; RT Am. J. Pathol. 158:617-625(2001).		CC -1- DR PROSITE; PS00019; ACTININ_1; FALSE_NEG.
RA		RA CC -1- FUNCTION: INTERLINKS INTERMEDIATE FILAMENTS WITH MICROTUBULES AND MICROFILAMENTS AND ANCHORS INTERMEDIATE FILAMENTS TO DESMOSOMES OR HEMIDESMOSOMES. COULD ALSO BIND MUSCLE PROTEINS SUCH AS ACTIN TO MEMBRANE COMPLEXES IN MUSCLE. MAY BE INVOLVED NOT ONLY IN THE CROSSLINKING AND STABILIZATION OF CYTOSKELETAL INTERMEDIATE FILAMENTS NETWORK, BUT ALSO IN THE REGULATION OF THEIR DYNAMICS.		CC -1- DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
RA		RA CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.		CC -1- DR PROSITE; PS00021; CH; 2.
RA		RA CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3, ARE PRODUCED BY ALTERNATIVE SPlicing.		CC -1- DR InterPro; IPR001589; Actinin_act_bind.
RA		RA CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in muscle, heart, placenta and spinal cord.		CC -1- DR InterPro; IPR001590; Actinin_act_bind.
RA		RA CC -1- DOMAIN: THE N-TERMINUS INTERACTS WITH ACTIN, THE C-TERMINUS WITH VIMENTIN, DESMIN, GFAP, CYTOKERATINS, LAMIN B; WHEREAS BOTH THE N- AND THE C-TERMINUS CAN BIND INTEGRIN BETA-4.		CC -1- DR InterPro; IPR001591; Actinin_act_bind.
RA		RA CC -1- PTM: PHOSPHORYLATED BY CDC2; REGULATES DISSOCIATION FROM		CC -1- DR InterPro; IPR001592; Actinin_act_bind.

FT	REPEAT	3975	4008	PLECTIN 21.	QY	195	KDEMILSLHYSNALQEKEELASRCRRLSQEELYLKQEL-.ORANMVSSCCELEQEQ-- 249		
FT	REPEAT	4063	4100	PLECTIN 22.	Db	2463	KAEAEIJQ-----QKELAQEQAERLQEDKEAQALAEETGFTLLEERQRQLEM 2515		
FT	REPEAT	4101	4138	PLECTIN 23.	QY	250	-----SLRTAS--DOESGDDEBLNLREEE---NEXLRSLTFSLAEBKDLDEQSLSDEAR 295		
FT	REPEAT	4139	4176	PLECTIN 24.	Db	2516	SAFAERKLKVAMSRAQARAEEADAQFERKQAEIGKLHRTELATEQKVTLVQTLELQR 2575		
FT	REPEAT	4177	4214	PLECTIN 25.	QY	296	GSRQELYVERITHSRLRERAAEAEROYEEKWEEKTQJQFQSKMCA---OLYREKVNALQ 351		
FT	REPEAT	4218	4252	PLECTIN 26.	Db	2576	QOSDHDER--IREAIAELEREKEKL--QQEARLQLQURESEEQTVQDQEQLQETQALQQ 2630		
FT	REPEAT	4265	4305	PLECTIN 27.	QY	352	AQYCE---LOKER-----DQAYSARDSQAIREISQSLVEREDSL----- 385		
FT	REPEAT	4319	4356	PLECTIN 28.	Db	2631	SFLSEKDSSLQRERFIEOEKALEQFLFODEVAQKQOREOORQOOQMEOERORLVASME 2690		
FT	REPEAT	4408	4445	PLECTIN 29.	QY	336	--RRQVIFELTDQVCELLTQLRQLOAEPPGVLKOEARTREPCCPREKQLVVRMHAICPRBDS 443		
FT	REPEAT	4446	4483	PLECTIN 30.	Db	2691	EARRRQHEAEEGVRRKQEELQOLE-----QORRQOEELAEEQRQLREQLQLEEQH 2742		
FT	REPEAT	4484	4521	PLECTIN 31.	QY	444	DCLVSS--TESQQLSDLSATSSRELVDSSFRSSSPAPPSSOOLSYKRVAYDGFGEPPWSFS 500		
FT	REPEAT	4522	4559	PLECTIN 32.	Db	2743	RANLAHSPEVTASQVAATKLPNGRDALDPAAEPHSDGLRRKVSAQRLQEAGLIS 2802		
FT	REPEAT	4560	4597	PLECTIN 33.	QY	501	SCLEIPEPDGPALGAKAGDPHDYELLTDALPQLESSLQPVSPGRLDVSESGVLMMR 560		
FT	DOMAIN	4250	4300	BINDING TO INTERMEDIATE FILAMENTS (BY SIMILARITY).	Db	2803	A-----EELORIAAOGHHTVVDE--LARRE 2823		
FT	DOMAIN	4625	4640	4 X 4 AA TANDEM REPEATS OF G-S-R-X.	QY	561	PARRILSVVTMLAFQGALLEQISVIGVGNLTFIHFVTPGSAAQDQMLRPGTQIVVADY 620		
FT	MOD_RES	4539	4539	PHOSPHORYLATION (BY CDC2) (BY SIMILARITY).	Db	2824	DYRHYL-----QGR-----SSINGLLKATNEKLSVYAAQLQRQLSPGTALILLEA 2869		
FT	VARSPLIC	1	174	MYAGMLMPRDQDQRAIYEVLFREGVYMAKKDRRPRPSLHPHPV GVINLQWNRMASTRAQVLRVLTNEGTAHL RQLHPLPEITPEQRLLRKKELEEVSETPTVVPATTORTLA GSPFKRGQPLPTPEQRLLRKKELEEVSETPTVVPATTORTLA REPOEPAPAT-SRSRGAGGAVSRAVSSDVSNSGGSSSPGD TLPWNLNGTORSRSRGAGGAVSLLPAERAVIRIA (IN ISOFORM 2, AND ISOFORM 3).	QY	621	EAS-----EPLFLAVLDTTLEAV--GLLRRVDGFCCLSVVKNTDGYK----- 662		
FT	VARSPLIC	409	412	MISSING (IN MD-ESS).	Db	2870	QAASGFLLDPVNRRL--TVNEAVKECVVGPFBLHHKULLSAEAVTGTGKDPTGQQTSLF 2926		
FT	VARIANT	429	429	L -> LL (IN MD-ESS).	QY	663	RULQD-----LEAKVATSG-----DSFYIYRNLMAMEGRAKGELOVHICNEVLHVT 706		
FT	VARIANT	1003	1005	/PTID=VAR_011336 MISSING (IN MD-ESS).	Db	2927	QANQKGILVREHNGIRLPLQIAITGGVTDVHSHRPDVAYR--RGYFDEEMNRVADP 2983		
FT	CONFLICT	185	185	/PTID=VAR_011337 S -> F (IN REF. 2).	QY	707	DTMFQGGCWHAHRVNSYTMKDTAAHGTIPNYNSRAQQQLIALIQ-MTQQC-TTVTRKPS 763		
FT	CONFLICT	259	259	N -> D (IN REF. 2).	Db	2984	SDDTKGFFDPNTHENLYL-----QLLERCYEDPPTGGLCULPLPLTKAA 3026		
FT	CONFLICT	550	550	H -> N (IN REF. 2).	QY	764	SGGPQKLVRSMDAKASPLRUSFDQRQLDP--SRMBSSTSFEWAECCLTVPTLYWP 821		
FT	CONFLICT	560	560	I -> V (IN REF. 2).	Db	3027	KGG-----ELVYT-----SEARDVFEKATVSAPFGKFOGKTVTYW-EIINSEYFT---- 3071		
FT	CONFLICT	706	706	Q -> R (IN REF. 2).	QY	822	HREPAPRPVPLYPRAVGLKCLQGFKKCLAEYLSQEEYEAWNSQGDLIQEG 877		
FT	CONFLICT	886	886	N -> Y (IN REF. 2).	Db	3072	--AEQRDILRORFRTGRITVEKI----KIIUTVVEEQE---OKGRLCFEG 3113		
FT	CONFLICT	1002	1002	V -> A (IN REF. 2).	RESULT	14			
FT	CONFLICT	1309	1309	V -> L (IN REF. 2).	ID	ZO3_MOUSE	STANDARD;	PRT;	905 AA.
FT	CONFLICT	1321	1321	L -> V (IN REF. 2).	AC	Q90XY1;			
FT	CONFLICT	1334	1334	V -> L (IN REF. 2).	DT	16-OCT-2001	(Rel. 40, Created)		
FT	CONFLICT	1534	1534	I -> M (IN REF. 2).	DT	16-OCT-2001	(Rel. 40, Last sequence update)		
FT	CONFLICT	1662	1662	A -> T (IN REF. 2).	DT	16-OCT-2001	(Rel. 40, Last annotation update)		
FT	CONFLICT	1688	1688	WLC -> RLR (IN REF. 2).	DE	Tight junction protein ZO-3 (Zonula occludens 3 protein) (Zona occludens 3 protein) (Tight junction protein 3).			
FT	CONFLICT	1767	1767	Q -> E (IN REF. 2).	DE	DE Occludens 3 protein (Tight junction protein 3).			
FT	CONFLICT	1789	1789	L -> A (IN REF. 2).	GN	Top3 OR ZO3.			
FT	CONFLICT	1910	1910	K -> R (IN REF. 2).	OS	Mus musculus (Mouse).			
FT	CONFLICT	2154	2154	K -> N (IN REF. 2).	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
FT	CONFLICT	2160	2160	S -> R (IN REF. 2).	OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.			
FT	CONFLICT	2215	2215	S -> Q (IN REF. 2).	ON	[11]			
FT	CONFLICT	2244	2244	S -> A (IN REF. 2).	RN	SEQUENCE FROM N.A.			
FT	CONFLICT	3027	3027	K -> E (IN REF. 2).	RP	STRAIN=DALBC;			
FT	CONFLICT	3310	3310	E -> A (IN REF. 2).	RC	MEDLINE=20056797; PubMed=10601346;			
FT	CONFLICT	3361	3361	F -> L (IN REF. 2).	RA	NCBI_TaxID=10090;			
FT	CONFLICT	3408	3408	F -> F (IN REF. 2).	RA	"Direct binding of three tight junction-associated MAGUKs, 20-1, 20-2,			
FT	CONFLICT	3447	3447	S -> A (IN REF. 2).	RT				
FT	CONFLICT	3531	3531	G -> A (IN REF. 2).					
FT	CONFLICT	3580	3580	R -> S (IN REF. 2).					
FT	CONFLICT	3589	3589	K -> Q (IN REF. 2).					
FT	CONFLICT	3596	3596	E -> Q (IN REF. 2).					
FT	CONFLICT	3616	3616	N -> H (IN REF. 2).					
FT	CONFLICT	3686	3686	V -> A (IN REF. 2).					
FT	CONFLICT	3786	3786	G -> A (IN REF. 2).					
FT	CONFLICT	3808	3808	K -> R (IN REF. 2).					
FT	CONFLICT	3816	3816	G -> A (IN REF. 2).					
Query Match			4.8%	Score 249; DB 1; Length 4684;					
Best Local Similarity			22.4%	Pred. No. 0.00015;					
Matches	187;	Conservative	112;	Mismatches 319; Indels 218; Gaps 34;					
Qy	135	LOBELNOQKQVAEEAARLSSVAQEAARLQIAEEDLAQQLAEMKLMKEKKM	194						
Db	2403	LOBEAEMKQVAEEAARLSSVAQEAARLQIAEEDLAQQLAEMKLMKEKKM	2462						

and ZO-3, with the COOH termini of claudins.";

J. Cell. Biol. 14:1351-1365 (1990).

-|- SUBUNIT: INTERACTS WITH OCCLUDIN, CLAUDINS AND ZO-1.

-|- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.

-|- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

-|- SIMILARITY: CONTAINS 1 KINASE-LIKE DOMAIN.

-|- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

C EMBL; AAF157006; AAF24175.1; -.
R HSSP; P31016; 1BFE.
R MGD; MGI:1351650; Tip3.
R InterPro; IPR000619; Guanylate_kin.
R InterPro; IPR01478; PDZ.
R InterPro; IPR01452; SH3.
R Pfam; PF00625; Guanylate_kin; 1.
R Pfam; PF00555; PDZ; 3.
R SMART; SM00072; GuKc; 1.
R SMART; SM00228; PDZ; 3.
R SMART; SM00326; SH3; 1.
R PROSITE; PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
R PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
R PROSITE; PS50102; SH3; FALSE_NEG.
R PROSITE; PS50106; PDZ; 3.
W Tight Junction; SH3 domain; Repeat; Membrane.
T DOMAIN 1 1 93 PDZ 1.
T DOMAIN 187 264 PDZ 2.
T DOMAIN 368 434 PDZ 3.
T DOMAIN 467 540 SH3.
T DOMAIN 654 754 GUANYLATE_KINASE.
SEQUENCE 905 AA: 99324 MW: B787BA1594661FEE CRC64;
Q

```

FX MEDLINE=92003925; PubMed=1912569;
 RA Toothaker L.E., Gonzalez D.A., Tung N., Lemons R.S., le Beau M.M.,
 RA Arnault M.A., Clayton L.K., Tenen D.G.;
 RT "Cellular myosin heavy chain in human leukocytes: isolation of 5' CDNA clones, characterization of the protein, chromosomal localization, and upregulation during myeloid differentiation.";
 RT Blood 78:1826-1833(1991).
 RL [3]
 RN RP SEQUENCE OF 1-715 FROM N.A.
 RX MEDLINE=91316803; PubMed=1860190;
 RA Simons M., Wang M., McBride O.W., Kawamoto S., Yanakawa K.,
 RA Gdula D., Adelsheim R.S., Weir L.;
 RT "Human nonmuscle myosin heavy chains are encoded by two genes located on different chromosomes.";
 RT Circ. Res. 69:530-539(1991).
 RN RP SEQUENCE OF 714-1960 FROM N.A.
 RX MEDLINE=90138950; PubMed=1967836;
 RA Saez C.G., Myers J.C., Shows T.B., Leinwand L.A.;
 RT "Human nonmuscle myosin heavy chain mRNA: generation of diversity through alternative polyadenylation.";
 RT Proc. Natl. Acad. Sci. U.S.A. 87:1164-1168 (1990).
 RN RP VARIANT DFN17 HIS-705.
 RX MEDLINE=20489856; PubMed=11023810;
 RA Lalwani A.K., Goldstein J.A., Kelley M.J., Luxford W., Castlein C.M.,
 RA Mhatre A.N.;
 RT "Human nonsyndromic hereditary deafness DFN17 is due to a mutation in nonmuscle myosin MYH9.";
 RL Am. J. Hum. Genet. 67:1121-1128(2000).
 RN RP VARIANTS MHA/FNNS/SBS K-93; C-702; C-1165; H-1424 AND K-1841.
 RX MEDLINE=20428192; PubMed=1097259;
 RA Seri M., Cusano M., Gangarossa S., Caridi G., Bordo D., Lo Nigro C.,
 RA Ghiglieri G.M., Ravazzolo R., Savino A., Del Vecchio M., d'Apolito M.,
 RA Iolascon A., Zelante L.L., Savoia A., Baldunini C.L., Norris P.,
 RA Magrini U., Belletti S., Heath K.E., Babcock M., Glucksmann M.J.,
 RA Aliprandis E., Bizzarro N., Desnick R.J., Martignetti J.A.,
 RT "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechner and Sebastian syndromes.";
 RT Nat. Genet. 26:103-105 (2000).
 RN [7] RP VARIANTS MHA ILE-1155 AND LYS-1841.
 RX MEDLINE=20428193; PubMed=1097260;
 RA Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.;
 RA Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.;
 RT "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in May-Hegglin anomaly";
 RT Nat. Genet. 26:106-10 (2000).
 CC -!- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS, CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND CAPPING.
 CC -!- DISEASE: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC) AND 2 LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- DOMAIN: THE RODLIE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 PEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF MAY-HEGGLIN ANOMALY (MHA), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.
 CC -!- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF FECHTER SYNDROME (FTNS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS, WITH ADDITIONAL ALPORT-LIKE CLINICAL FEATURES OF SENSORINEURAL DEAFNESS, CATARACTS AND NEPHRITIS.
 CC -!- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF SEBASTIAN SYNDROME (SBS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.
 CC -!- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF AN AUTOSOMAL DOMINANT FORM OF NONSYNDROMIC SENSORINEURAL DEAFNESS (DFNA17) WHICH IS CHARACTERIZED BY PROGRESSIVE HEARING IMPAIRMENT AND COCHLEOSACCULAR DEGENERATION.
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR EMBL: Z2215; CAB05105_1; -.
 CC DR EMBL: M81105; AAA59888_1; -.
 CC DR EMBL: M69180; AAA61785_1; -.
 CC DR EMBL: M31013; ARA36349_1; -.
 CC DR HSPP: P08759; ILVK.
 CC DR MM: 160775; -.
 CC DR MM: 153640; -.
 CC DR MM: 155100; -.
 CC DR MM: 603622; -.
 CC DR MM: 605249; -.
 CC DR InterPro: IPR000048; IQ.
 CC DR InterPro: IPR000009; Myosin_N.
 CC DR InterPro: IPR0032928; Myosin_tail.
 CC DR InterPro: IPR002017; Spectrin.
 CC DR InterPro: IPR001609; myosin_head.
 CC DR PF00512; IQ; 1.
 CC DR PF00063; myosin_head; 1.
 CC DR PF02736; Myosin_N; 1.
 CC DR PF01576; Myosin_tail; 1.
 CC DR PRINTS: PRO0193; MYOSINHEAVY.
 CC DR PRODOM: PD000355; myosin_head; 1.
 CC DR SMART: SM00015; IQ; 1.
 CC DR SMART: SM00242; MYSC; 1.
 CC DR PROSITE: PS50096; IQ; 1.
 CC KW Myosin; ATP binding; Calmodulin-binding; Actin-binding;
 CC KW Coiled coil; Alkylation; Multigene family; Disease mutation;
 CC KW Deafness.
 CC FT DOMAIN 1 778
 CC FT DOMAIN 779 808
 CC FT DOMAIN 837 1926
 CC FT DOMAIN 837 181
 CC FT BIND 174 181
 CC FT DOMAIN 654 676
 CC FT MOD_RES 694 694
 CC FT MOD_RES 704 704
 CC FT VARIANT 93 93
 CC FT VARIANT 1155 1155
 CC FT VARIANT 702 702
 CC FT VARIANT 1165 1165
 CC FT VARIANT 705 705
 CC FT VARIANT 1424 1424
 CC FT VARIANT 1841 1841
 CC FT CONFLICT 53 55
 CC FT CONFLICT 660 660
 CC FT CONFLICT 869 869
 CC FT CONFLICT 931 931
 CC FT CONFLICT 1240 1241
 CC FT CONFLICT 1350 1350
 CC FT CONFLICT 1764 1754
 CC FT CONFLICT 1771 1771
 CC SQ SEQUENCE 1960 AA; 226531 MW;

Query Match 4.7% Score 240.5; DB 1; Length 1960;
 Best Local Similarity 21.1%; Pred. No. 0.00014; Mismatches 359; Indels 353; Gaps 59;
 Matches 238; Conservative 176;

QY 14 LDEBTWMMESHRH--RIVRCICPS----RLTYLRAKVLCOLDEEVLYHSPR-- 62
 Dd 649 LYKEQLAKLMATLRTNTNPNEVRCIIPHEKKGLDPLH--VLDOLRCNGVLEGRIC 704
 QY 63 -----LTHSAMRAGHLJUDLLKTRKGNGAIAFLESLSKPHNPDVYT- 102
 Dd 705 RGFPNRVVFQFRQRVEILTNSIPKGM----DGKQACVLMKALELDSNLRYIG 757
 QY 103 -----VNGQPDVDFNSFSGLMETSKUITECLAG-----AIGSLOEELN 140
 Dd 758 QSRVEFRAGVLHLHEEFRDL-----KITDVLLGFAQCRGYLARKAFAKRQQQLT 807
 QY 141 QEKQKEVYLRLRCQQLBEHILGAE-----TRAEGLHQLEADHSRMKREVSAYHFHEVLR 193
 Dd 808 ANK---VLRNC--AYALKLRNQWFLRFKVKPILQV---SRQEEMMAKEEVLK 856
 QY 194 LKDEMLSLHYSNALQEKELLASCRCSRSLQEFYLKLQELQRANMYSCELELQEQLR- 252
 Dd 857 VREKQLA---ENRLTEME----TLQSLMLAEFLQJB-QLGAETELCAEBELRA 904
 Qy 253 -TASDOEGSDPEL----NRLKEENEKLRSLTFSLAEKDILEQSDEARSRQELVERI 305
 Dd 905 RUTAKKBL--EJICHDLEARVEEEERQHQI-L--AEKKMQNQIQLQEEPEEESAR 959
 Qy 306 HSLRERAVAAEFOREQWEEKETOLLOFQKSXMACOLYRKEVNLAQOVC---LQKER 361
 Dd 960 QRLQLEKVTTERPLKLL--EBEQILEDQN---CKLAKER-KLLEDRIAEFTTNUTEE 1012
 Qy 362 DAYS-----ARDSAQR--EISQSLYEDS-LRROYFELTDQVC 397
 Dd 1013 EKSKSLAKLKKNHEAMITDLEERLRRREEKQROLEKTRRKLEGSTDLSQIAELAQIA 1072
 Qy 398 ELRTOLRQLQAEPGPGVLKQEARTRPCEPKRQLVYRHMACPDDSDCLSYSTESOLLS 457
 Dd 1073 EKMQMLAKKEELQQAII--ARVEEEAQKNAKHLKIREL-----ESQ-IS 1114
 Qy 458 DLSATSSRELVDSSRSSSPAPPSQSQSLYKRVADEFGEEPWSSCCLIEPGDPGALPGAK 517
 Dd 1115 EL---QEDUESERASRNKAEKQ-----RDIGEE-----LEA----- 1143
 Qy 518 AGDPHDYELLETTADIPQLESSLQYPSPGRDVSSEG-----LMRRPARRI 565
 Dd 1144 ---LKTELELTLDSTAQAQQLRSKREQEYNILKTLLEEAFTKTHEAQIQEMRQHSQ-- 1196
 Qy 566 LSQVTMIAFGQDALLRQDRIOSVIGGNL-----TGFTHR-----TPGSAADQMAL 609
 Dd 1197 -AVEELAEQ---LEOTKRYKANLEKAKOTLENERGELANEKVILLOGKDSEHKKKV 1250
 Qy 610 RFTGOLVMYDYEASEPLFKAVIEDTT-----LEEAVGLLRRYDGFCCLSYKNTD-GYK 662
 Dd 1251 EAQLQELQVKFNEGERVTELAJKVTKLQYLDNVTGLLSODS--KSSRLTKDFSALE 1307
 Qy 663 RLQDLE-----AKVATSGDSFYIRVNLAMEGRAKGELO---VHCN 700
 Dd 1308 SQLODQELLIQEENRQLSLSTKLQVEDENSP--REQEEEEEAKHNLJEKQIATLHA- 1364
 Qy 701 EVLHVITDTMFOCGCMH-AHNVNSTMKD-----TAHGTI-PNSRAQQQLIA 747
 Dd 1365 QADMKKKMDMSVGCLEAEEVKRLQDKLEGSLQRHEEKVAYOKLEKTKTRLGQELDD 1424
 Qy 748 LJDQMTQQ----CTVTRKPSSGGPQLVRVSMDKAKASPLRLSPDRQOLDPSRMEGSST 803
 Dd 1425 LLVLDLHQRSACNLXK---QKFDQDQIAEKTISAKYAEERDAAERETKA- 1478
 Qy 804 CFWAESCLTLVPTYLWPHFRPARPREVLLYRVAYGKILSEKCLLQGFKKLAELSQEE 863
 Dd 1479 -----LSLARALBEAMEOKAL---ERLNKQFRTEM 1506
 Qy 864 YEAWSORQDI--IQGEVSGRCWVTRHASESLMKNTHALDVSCLHMDIFP 920
 Dd 1507 EDLMESSKDDYCKSVBIEKS-----KRALEQQYEMKTMQLEELDEDQATEDAK--- 1555
 Qy 921 IVIHVSVN-EKMAKLUKGLO-RUGTSEE--QILLEAROEQEGDLD 961

QY	532	DLPQELESSIQWPSP--GRUDVS--ESGVLMRRRPARRILSQTMLAFOGDALEQIISVIG	587	QY	542	PVSPGR-----LDV-----SESVLW-----RRPARRILSQTMLAFOGDAALL	580
Db	400	- - - - -PSPSDPAIPLNPSNAHEDGTL--RPSMPLKLVK-- - - - -FRKGSV - GURLAG	439	Db	358	PVTPTRGRSSAAIDEPPRPPPROSSGGAAQDEFYSSRQLYBEROASPREFSFQEKGSV	417
QY	588	GNLTGIFLHRVPCGSAADOMALRPGTQIVMVDEASEPLFKAVLDTTL-- - - - -BEA	639	QY	581	EQISVIGGNLTGIFLHRVPCGSAADOMALRPGTQIVMVDEASEPLFKAVLDTTL	640
Db	440	GNDVGIFAYGVLDFSPAKEGLEFGDQLRVRN-- - - - -NVDFTNLIREEAVFLDLPKGEEV	496	Db	418	-GIRLGNEAGIFTAVQPGSPASLQLMPGDKILKVN-----DMDMNGVTREAV	468
QY	640	VGLLRVVDGFCCLSVKVNTDGKYLRLQDIEAKVATSGSDSFYIRVNLAMEGRAGELOQHC	699	QY	641	GUJLRRVDFCCLSVKVNTDGKYLRLQDIEAKVATSGDSFYIRVNLAMEGRAGELOQHV	698
Db	497	TIAQK-----KKDVKYRIVE-----SDVGDSFYIRTHFEYEKEPSYGLSFK	539	QY	699	CNEVLYHDTMFMQ-CGCMWAHRVNSYTMKDTAAHTGIPNYSRAQQQLIALFQDMTOQCT	757
QY	700	NEYLHVTDTMFQ-CGCMWAHRVNSYTMKDTAAHTGIPNYSRAQQQLIALFQDMTOQCT	758	Db	469	LFL-----LSLQDRDLTIVYCKEYDEVTNQRGDSFHKTHPHCDNFSKGEMAFK	520
Db	540	GEYERVVDTLYNGRLGSWAIRIGKHNHE--VERGILPKRNRAEQ -LASVQ-----Y	588	Db	521	AGDVERVDTLYANGVVGWQVLIGRQHE--MORGVLPNKSAAEL-----ATAQFN	571
QY	759	TRKPSSGGPKQLYRIVSMSDKAKASPLR-LSFDRGQLDPSMGSSTCEWAESCLTLVP-Y	816	QY	758	VTRKPSSGGPKQLYRIVSMSDKAKASPLR-----NEGSTTCFWEASCL	811
Db	589	TPKTAGG-----DRAFWRFGRGLRSRKRSLRSLDLS-----AQPVQTKPBY	633	Db	572	ATTK-----EMMANESGNCNFERRRSTHRSKSLSRNWWDDYVFSDIS	615
QY	817	TLYWPHRPARPAPRPPVLLYPPRAVGKTLSEKUCLLGFKKCLLAETLSQEE--YE AWSQRGD	872	QY	812	TLYWPHRPARPAPRPPVLLYPPRAVGKTLSEKUCLLGFKKCLLAETLSQEE	871
Db	634	ERVYLREAGFLRPVTF-GPIADYARE-----LAREEPDIYQTAKSELRD	678	Db	616	KFPAYERVYLRHGFREVYVLF-GPVSPSLARELIA-KDFDPDKFSTPLQDDKSA	667
QY	873	IIOGEVSGGRGCWVTRHAEVSMKEHTHALDVSCTYLTMRHDIFPVIHVSVNERMA	932	QY	872	DIIQEGEVSGGRGCWVTRHAEVSMKEHTHALDVSCTYLTMRHDIFPVIHVSVNEK	930
Db	679	AGTDHRSSG--TIRLHTKIQIDQKHALLDVTNPVNADRLNQAQWPIVFLNPDSKQG	735	Db	668	--ATSGKCRIVRSLNRDVMRKHALLDVTNPVNADRLNQAQWPIVFLKTISK	720
QY	933	KKL-----KKGLQRLGTSEEQLEAARQBEQGDLDRAPCLYSSLAPDGWS-----SDL	983	QY	931	-MAKKLKKGQLRG-TSEQLLEAARQBEQGDLDRAPCLYSSLAPDGW-----SDL	978
Db	736	VKTMRMRLCPESRKSKARKYERSHKL---RKNNHHLFTTTINLNSM-NDGW-----YG	784	Db	721	HVTKQLRIGLKPDAHKSKRKLLEBQCK-----LBRV-----WSHIFTQIALSDE	765
QY	984	CVROTAIADEQKVYVTEQ	1001	QY	979	DGLLSCVROTAIADEQKVYVTEQ	1002
Db	785	ALKEAIQQQNQLVWVSE	802	Db	766	ESWYRKLRDSLIDLQSGAVWMSSE	789
RESULT 2							
QY	13703	tamou protein - fruit fly (<i>Drosophila melanogaster</i>)		A47747	tight junction protein 20-1 - human		
C;Species:	<i>Drosophila melanogaster</i>		C;Species: <i>Homo sapiens</i> (man)	C;Cross-references:	GB:114937; NID:9292937; PID:AAA02891_1; PID:g292938		
C;Date:	13-Aug-1999	#sequence_revision 13-Aug-1999	C;Date: 02-Jun-1994 #text_change 05-Nov-1999	C;Accession:	A47747		
C;Accession:	T13703		R;Willcott, E.; Balda, M.S.; Fanning, A.S.; Jameson, B.; Van Itallie, C.; Anderson, J.	R;Title:	The tight junction protein 20-1 is homologous to the <i>Drosophila</i> discs-large protein. The <i>Drosophila</i> tamou gene, a component of the activating pathway of extramacroc		
R;Takahashi, M.; Togashi, S.; Suzuki, T.; Kobayashi, M.; Murayama, A.; Kondo, K.; Miyake, Genes Dev 10, 1781-1795, 1996	A;Reference number:	Z11700; MUID:96312452	A;Reference number:	A47747; MUID:93361541	A;Accession:	F:15-94/Domain: GLGF domain homology <GLG>	
A;Accession:	T13703	A;Status: preliminary; translated from GB/EMBL/DDJB	A;Status: preliminary	A;Accession:	F:16-486/Domain: GLGF domain homology <GLG>		
A;Molecule type: mRNA	A;Molecule type: mRNA	A;Residues: 1-1367 <TAK>	A;Residues: 1-1736 <WIL>	A;Molecule type: mRNA	F:632-782/Domain: guanylate kinase homology <GLG3>		
A;Cross-references:	EMBL:D83477; PID:91498136; PID:BN11923_1;	A;Cross-references:	A;Cross-references: GB:114937; NID:9292937; PID:AAA02891_1; PID:g292938	C;Superfamily:	guanylate kinase homology; GLGF domain homology		
C;Genetics:		C;Keywords:	C;Keywords: alternative splicing; membrane protein; phosphoprotein	F:15-94/Domain:	F:151-248/Domain: GLGF domain homology <GLG>		
A;Gene: tamou (tam)		A;Gene: tamou (tam)	A;Gene: tamou (tam)	A;Accession:	F:416-486/Domain: GLGF domain homology <GLG>		
A;Cross-references:	FlyBase:FBgn0003177			A;Residues:	F:632-782/Domain: guanylate kinase homology <GLG>		
Query Match 2							
QY	373	REISOSLVEENDSLRQFEDTQVCELRQLRQLOQAEPGPVQLQEARIPECPREKORLV	432	Query Match 2	Score 270.5; DB 2; Length 1736;		
Db	194	KEISSK- ⁷ AREQLNANGYS ¹¹ - ¹¹ Q ¹² DI ¹³ TRI ¹⁴ NTNCGDTMSLKEAKKIDGC-KERI ¹⁹ EV	248	Best Local Similarity 21.4%; Pred. No. 2.4e-06;			
QY	433	RMHAICPRD-SDCSLVSSTESTOLSLDSAT-----SSRELVDSSRSSPAPPSSQ	481	Matches 174; Conservative 129; Mismatches 278; Indels 233; Gaps 38;			
Db	249	VLRTDTNQAVSQLNNSASHQASQNSYIATHOPQVSGCSSSNLNLEDPLPG-ASVSS	307	Qy 247 QEQSLRATSDQESDEE-----UNRLKEENKRSLSLAEDILEQSDLRG	296		
QY	482	QSLYKRVADFEEPWFSQSSCLEPEGDGAFLAKADPHDYLDTADPOLESSQ	541	Db 151 RERSLSPDRRSVASSQPAKPTKVTLKSRKNEYEGLR-LASHIFVKEISQDSLAA	209		
Db	308	QNLVQPPTRTSNGPNINGNLN - DEKSNLTPRGRSRGP-----IMCGVSQQLD--R	357	Qy 297 SROK-EYERIH-SLERAVAEEQRECYEEKTLQFOKSMACYREKVNALQAO	353		
Query Match 3							
QY	13703	tamou protein - fruit fly (<i>Drosophila melanogaster</i>)		Qy 632-782/Domain: guanylate kinase homology <GLG>			
C;Species:	<i>Drosophila melanogaster</i>		C;Species: Homo sapiens (man)	C;Cross-references:	GB:114937; NID:9292937; PID:AAA02891_1; PID:g292938		
C;Date:	13-Aug-1999	#sequence_revision 13-Aug-1999	C;Date: 02-Jun-1994 #text_change 05-Nov-1999	C;Accession:	A47747		
C;Accession:	T13703		R;Willcott, E.; Balda, M.S.; Fanning, A.S.; Jameson, B.; Van Itallie, C.; Anderson, J.	R;Title:	The tight junction protein 20-1 is homologous to the <i>Drosophila</i> discs-large protein. The <i>Drosophila</i> tamou gene, a component of the activating pathway of extramacroc		
R;Takahashi, M.; Togashi, S.; Suzuki, T.; Kobayashi, M.; Murayama, A.; Kondo, K.; Miyake, Genes Dev 10, 1781-1795, 1996	A;Reference number:	Z11700; MUID:96312452	A;Reference number:	A47747; MUID:93361541	A;Accession:	F:15-94/Domain: GLGF domain homology <GLG>	
A;Accession:	T13703	A;Status: preliminary; translated from GB/EMBL/DDJB	A;Status: preliminary	A;Accession:	F:416-486/Domain: GLGF domain homology <GLG>		
A;Molecule type: mRNA	A;Molecule type: mRNA	A;Residues: 1-1367 <TAK>	A;Residues: 1-1736 <WIL>	A;Residues:	F:632-782/Domain: guanylate kinase homology <GLG>		
A;Cross-references:	EMBL:D83477; PID:91498136; PID:BN11923_1;	A;Cross-references:	A;Cross-references: GB:114937; NID:9292937; PID:AAA02891_1; PID:g292938	C;Superfamily:	guanylate kinase homology; GLGF domain homology		
C;Genetics:		C;Keywords:	C;Keywords: alternative splicing; membrane protein; phosphoprotein	F:15-94/Domain:	F:151-248/Domain: GLGF domain homology <GLG>		
A;Gene: tamou (tam)		A;Gene: tamou (tam)	A;Gene: tamou (tam)	A;Accession:	F:416-486/Domain: GLGF domain homology <GLG>		
A;Cross-references:	FlyBase:FBgn0003177			A;Residues:	F:632-782/Domain: guanylate kinase homology <GLG>		
Query Match 3							
QY	373	REISOSLVEENDSLRQFEDTQVCELRQLRQLOQAEPGPVQLQEARIPECPREKORLV	432	Query Match 3	Score 270.5; DB 2; Length 1736;		
Db	194	KEISSK- ⁷ AREQLNANGYS ¹¹ - ¹¹ Q ¹² DI ¹³ TRI ¹⁴ NTNCGDTMSLKEAKKIDGC-KERI ¹⁹ EV	248	Best Local Similarity 21.4%; Pred. No. 2.4e-06;			
QY	433	RMHAICPRD-SDCSLVSSTESTOLSLDSAT-----SSRELVDSSRSSPAPPSSQ	481	Matches 174; Conservative 129; Mismatches 278; Indels 233; Gaps 38;			
Db	249	VLRTDTNQAVSQLNNSASHQASQNSYIATHOPQVSGCSSSNLNLEDPLPG-ASVSS	307	Qy 247 QEQSLRATSDQESDEE-----UNRLKEENKRSLSLAEDILEQSDLRG	296		
QY	482	QSLYKRVADFEEPWFSQSSCLEPEGDGAFLAKADPHDYLDTADPOLESSQ	541	Db 151 RERSLSPDRRSVASSQPAKPTKVTLKSRKNEYEGLR-LASHIFVKEISQDSLAA	209		
Db	308	QNLVQPPTRTSNGPNINGNLN - DEKSNLTPRGRSRGP-----IMCGVSQQLD--R	357	Qy 297 SROK-EYERIH-SLERAVAEEQRECYEEKTLQFOKSMACYREKVNALQAO	353		
Query Match 4							
QY	13703	tamou protein - fruit fly (<i>Drosophila melanogaster</i>)		Qy 632-782/Domain: guanylate kinase homology <GLG>			
C;Species:	<i>Drosophila melanogaster</i>		C;Species: Homo sapiens (man)	C;Cross-references:	GB:114937; NID:9292937; PID:AAA02891_1; PID:g292938		
C;Date:	13-Aug-1999	#sequence_revision 13-Aug-1999	C;Date: 02-Jun-1994 #text_change 05-Nov-1999	C;Accession:	A47747		
C;Accession:	T13703		R;Willcott, E.; Balda, M.S.; Fanning, A.S.; Jameson, B.; Van Itallie, C.; Anderson, J.	R;Title:	The tight junction protein 20-1 is homologous to the <i>Drosophila</i> discs-large protein. The <i>Drosophila</i> tamou gene, a component of the activating pathway of extramacroc		
R;Takahashi, M.; Togashi, S.; Suzuki, T.; Kobayashi, M.; Murayama, A.; Kondo, K.; Miyake, Genes Dev 10, 1781-1795, 1996	A;Reference number:	Z11700; MUID:96312452	A;Reference number:	A47747; MUID:93361541	A;Accession:	F:15-94/Domain: GLGF domain homology <GLG>	
A;Accession:	T13703	A;Status: preliminary; translated from GB/EMBL/DDJB	A;Status: preliminary	A;Accession:	F:416-486/Domain: GLGF domain homology <GLG>		
A;Molecule type: mRNA	A;Molecule type: mRNA	A;Residues: 1-1367 <TAK>	A;Residues: 1-1736 <WIL>	A;Residues:	F:632-782/Domain: guanylate kinase homology <GLG>		
A;Cross-references:	EMBL:D83477; PID:91498136; PID:BN11923_1;	A;Cross-references:	A;Cross-references: GB:114937; NID:9292937; PID:AAA02891_1; PID:g292938	C;Superfamily:	guanylate kinase homology; GLGF domain homology		
C;Genetics:		C;Keywords:	C;Keywords: alternative splicing; membrane protein; phosphoprotein	F:15-94/Domain:	F:151-248/Domain: GLGF domain homology <GLG>		
A;Gene: tamou (tam)		A;Gene: tamou (tam)	A;Gene: tamou (tam)	A;Accession:	F:416-486/Domain: GLGF domain homology <GLG>		
A;Cross-references:	FlyBase:FBgn0003177			A;Residues:	F:632-782/Domain: guanylate kinase homology <GLG>		

QY	593	IPTHRVTPGSAADQMLRPGTQIVDYESEPLEPKAVLDTITLEAVAGLLRVIDGFCCFL	652	QY	386	-RRQVPELTDOYCELRQLRQLQAEPPGVVLQEARIIPRECPREKQRVLMRAICPRDS	443
Db	807	--VKEYAAWRDYE--	-LM	Db	2581	EARRRQEAEEGVRRQEELEQOLE-----QQRQQEELAEEENQRLEQOLLEEQH	2632
QY	653	SVKYNTDGKYLRLQDLEAKVATSGDSFYIRVNNAMEGRAKLQVHCNEVLMHTDMFQG	712	QY	444	DCLSLLV---TESOLSLDSATSSRELVDSSFRSSPAFQSOSLYKRAEDGEFWPSFS	500
Db	837	TKECECKARQELQFAKEKA-GLESHSELQIISQQNKLAE-	-HAN	Db	2633	RAALANSEEVTSQVAATKTLPNGRDALGPAAEAEPHSFGDLRKVSAQRLOEAGILS	2692
QY	713	CGCWHAHRYNSYTMRKDTAAHTIPNYSRAQQQLIAQMTQOCVTVRKPSSGGPQKLV-	771	QY	501	SCLEIPPEGDPGALPGAKAGDPHYDYLDTADLPQLESSLQPVSPGRDVSSEGVLMRR	560
Db	881	--LARALQOQKEKEVRAOKLADDLSLQEJKMATSKEVARLEFLVRK-AGEQETAS	934	Db	2693	A-----EFLQRLAQGHITTVDE--LARRE	2713
QY	772	RIVSHDKAKASPLRLSFDQLDPSRM--GSSTCFWAESCCLTLYPTLVWPHRPAPR	828	QY	561	PARRLSQVTMIAFGDALLQISVIGGNLTGIFIHRTVPGSADOMMLRPGTQIVWVDY	620
Db	935	RELXREPARA-----GDRQPWNLLEQQGQRC-----STQA	965	Db	2714	DVRHYL----OGR-----SSLAGKLKATNEKLSSVYAAQLRQLSPSTALILEA	2759
QY	829	PVLLPVRAVKGKILSEKLCLLQFKKCLAEYLQSEEYEAWNSQRQDLTQEVEGSSGRCWTR	888	QY	621	EAS-----EPELKFAVLDTTLBAV-GLLRRVDFGCLLSVKVNTDGK-----	662
Db	966	ALQAMEREAFQMGNE---LERLRALMEQGQQOEERGQEREVARLTOERG-----R	1015	Db	2760	QAASGFLLPVNRRRL--TNEAVKGGVVGPELHHKILSAAERAVIGKYPGQIQISLF	2816
QY	889	HAVEISLMEKNTHALLDVQDLSVCTLHRMDIFPI--VIHVSXNEKMAK---KLR-----	936	QY	663	RLLQD-----LEAKVATSG---DSFYIRVNLAMEGRAKGELQVHCNEVLYHVT	706
Db	1016	AQADALEAKARAELMRQLNALNEQRFEPATQELAHALTEKEGKDQELAKLKGLEAA	1075	Db	2817	QAMQKGLTIVREHGRGIRLLDEAQIATGGVDPVHSRVPDVAYR--RSYFDEENNRVLA	2873
QY	937	--TKGLQRLGTSSEQUELEAAQEE	957	QY	707	DTMFQGCCWIAHRVNSYTMKDTAAHTIPNSRAQOQLIJD-MTQQC-TTVTRKPS	763
Db	1076	QIKEELBLROTQVQKLEQLAKE	1098	Db	2874	SDDTKGFFDPNTHENYL-----OLLERCVEDPETGCLCLPLTRKAA	2916
RESULT	5			QY	764	SGGPQKLVRIYSMDKAKASPLRUSFDRGQLDP-SRMEGSSTCFWAESCLTVPTLVWP	821
G02520		pectin - human		Db	2917	EGG-----SEARDVFERATVSAFPGKQGKTVTIW-EILINSEYFT----	2961
C	Species: Homo sapiens (man)			QY	822	HRPAPRPPVLLPRAVGKILSERKLCLQGFKRLAELYLSQEEYEAWNSQRGDIQEG	877
C;Accession: G02520	Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 # text_change 20-Aug-1999			Db	2962	--AEQRDRLLRQFRGTRITVEKII---KLIITVVEQE---QKGRLCFEG	3003
R;McLean, W.H.L.; Smith, F.J.D.				RESULT	6		
submitted to the EMBL Data Library, March 1996				A;Reference number: H01385			
A;Accession: G02520				A;Status: preliminary			
A;Molecule type: mRNA				A;Species: Homo sapien (man)			
A;Residues: 1-4574 <MCL>				C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002			
C;Cross-references: EMBL:U53204; NID:91477645; PID:91477646				R;Liu, C.G.; Maercker, C.; Castanon, M.J.; Hautermann, R.; Wiche, G.			
C;Genetics:				proc. Natl. Acad. Sci. U.S.A. 93, 4278-4283, 1996			
A;Gene: PLEC1				A;Title: Human plectin: organization of the gene, sequence analysis, and chronic			
C;Superfamily: Plectin; alpha-actinin actin-binding domain homology <ACT>				A;Reference number: C59404; MUID:96210632			
F;68-283/Domain: alpha-actinin actin-binding domain homology				A;Accession: C59404			
submitted to the EMBL Data Library, March 1996				A;Status: preliminary			
A;Molecule type: DNA				A;Molecule type: DNA			
A;Residues: 1-4684 <STO>				A;Cross-references: GB:CAA91196; PID:NID:91296662; PIDN:CAA91196.1			
C;Superfamily: Plectin; alpha-actinin actin-binding domain homology; ribosomal				C;Superfamily: Plectin; alpha-actinin actin-binding domain homology; ribosomal			
Query Match	4.8%	Score 249; DB 2; Length 4574;		Query Match	4.8%	Score 249; DB 2; Length 4684;	
Best Local Similarity	22.4%	Pred. No. 0.00012; Mismatches 319; Indels 218; Gaps 34		Best Local Similarity	22.4%	Pred. No. 0.00012; Mismatches 319; Indels 218; Gaps 34	
Matches 187; Conservative 112; Mismatches 319; Indels 218; Gaps 34;				Matches 187; Conservative 112; Mismatches 319; Indels 218; Gaps 34			
Db	2293	LQEEAEKMQVAEEAARLSTAQAEEAARLQAAEELAQRQALAEMLKERMQVAEATRL	2352	Qy	135	LQEEAEKMQVAEEAARLSTAQAEEAARLQAAEELAQRQALAEMLKERMQVAEATRL	194
Qy	195	KDEMQLSHSYNALSQKEELAASCRSLSQPEELYLKQEL-QRANNVSSCELELQE-----	249	Db	2403	LQEEAEKMQVAEEAARLSTAQAEEAARLQAAEELAQRQALAEMLKERMQVAEATRL	2462
Db	2353	KAEEELLQ-----QQKELAQEQARRLQDKEQMAQQLAEETQGQFTILEERQRLEM	2405	Qy	196	KDEMQLSHSYNALSQKEELAASCRSLSQPEELYLKQEL-QRANNVSSCELELQE-----	249
Qy	250	--SLRTAS---DQESGDEELNLKKE---NEKLRLSITFLSAAEKLQDLEQSLDEAR	295	Db	2463	KAAEELLQ-----QQKELAQEQARRLQDKEQMAQQLAEETQGQFTILEERQRLEM	2515
Db	2406	SEAERLKLKEVAAEMRAQAAEEDQFRKQAAEIEGEKHLRTELAQEKVTLVQPLEIQR	2465	Qy	250	--SLRTAS---DQESGDEELNLKKE---NEKLRLSITFLSAAEKLQDLEQSLDEAR	295
Qy	296	GSRQELVERITHSLRERAVAEEERQEOYEEKTQLQFQSKMAC---QLYREKVNALQ	351	Db	2516	SEAERLKLKEVAAEMRAQAAEEDQFRKQAAEIEGEKHLRTELAQEKVTLVQPLEIQR	2575
Db	2466	QDSHDAAER--LREIAEELERKEKL---QOEAKULLQKSEEMOTVQEQQLLQETQALQQ	2520	Qy	296	GSRQELVERITHSLRERAVAEEERQEOYEEKTQLQFQSKMAC---QLYREKVNALQ	351
Qy	352	AQVCE---LQER-----DQAYSARDSAQRTSLSVKEIDS-----	385	Db			
Qy	2521	SFLSERKDSLQQRERIEQEKAKLQFQDEVAKQQLREFQQQQMEOERQLYVASME	2580	Qy			

Db	2576	QOSDHAER--LREATAELEREKEKL--QQEAKUJLQLKSSEMQTVQEQLQETQALQQ	2630	Qy	229	LKQE-----LQRANMVSSCELELQFQSLRPTASDQESGDEBLNRKKEENEKLRSLTFSL	281
Oy	352	AGYCE---LORER-----DQAYSARDSAQREISQSLYVKDSL----	385	Db	559	LRQEOTIVTAALARABO-SIAIRSLSSNTLT-----EVADIRAAAYKLSAINEAL	608
Db	2631	SFLSEKDSLIGRERFIEQEKKAKLQLFQDEYAKAQOLREEQQRQQMEEQRQLVASYME	2690	Qy	282	A-EKDILEQSL----DEARSQELIVERI-HSLRRAVAAERQEQYWEKEQTILQF	333
Oy	386	-RROYFELTDQCEULRQLRQLAQEPGVLKTRPQPREKRLYRMAICPRDDS	443	Db	609	ALDKVGLINQQLIQLQEEENQSYCSRMPEAEQARNALQWDLAEEKRNEALWKNTLHQAL	668
Db	2691	EARRROHEAEBSVRKQEEQELQLE-----QQRQQEEELAEEQNLRLRQOLLEEQH	2742	Qy	334	QKSXMACOLYRERKVNALQAQYCELQERDQAYSARDASQAREISQSLSVEKDSSLRQYELT	393
Qy	444	DCSLLVS--TESOLLSLSATSSRSRELVDSTRSSPAPPSSQSLSYKRAEDEGEPWSFS	500	Db	669	QKAE-----EAGAEIQAQDLRDIKEKE-----IQKLKISESRHQSEA-----	705
Db	2743	RAALAHSEEVYASQVAATKTLUPNGRDALGPAEAEPHSDFGLRRKVSQARLQEAGILS	2802	Qy	394	DQYCELFITQRLQAE--PPGYLKQFOARTREPCTPREKFRL-VRMHAIICPRDDSDCSLVY	449
Oy	501	SCLEIPPGDPBALPGAKAGDHLDYELLDIAADLPQLESSLQPVSPRLDYESGVLMRR	560	Db	706	-----ATTQLEBLHQKEKQROQEEVLRARQKEEALVRKEAALVRLQAV-ERDQDL-----	755
Db	2803	A-----EELCRLAQCHTTDE--LARRE	2823	Qy	450	STESQILSLSATSSRSRLVDSFRSSSAPPSSQSLYKRAEDEGEEWFSSSCLEIPEGD	509
Oy	561	PARRILSQVTMLAFQGDALLEQISVIGGNLTGIFIFHRVTPGSAADMALRPGTQIVMVYD	620	Db	756	AAQQLGILS-----SAKELLES-----SLF-----EPQQQNSVIDEEQG-	788
Db	2824	DYRHYL-----QGR-----SSTAGLLIKATNEKLSSVAAQOLQLLSGTALLAE	2869	Qy	510	PGALPGAKAGDPHDYFLDFTADLPQLESSLQPVSPR-----LDVSESGVL	556
Qy	621	EAS-----EPLFKAYLEDTTLEFAV--GLLRFDGFCCLSVWKVNTDGYK-----	662	Db	789	-----QLEVQIOTVTQAKEVIOGEVRLKLELDTERSQA	823
Db	2870	QAAASGLFLDPVRNRRL--TVNEAVKEGVYGPHELHKLSSERA/TVGYKDPTGQOISL	2926	Qy	557	MRRRPARRILSQVTMLAFQGDALLEQISVIGGNLTGIFIHRVTPGSADOMALRPGTQIV	616
Oy	663	RLLQD-----LEAKYATSG-----DSFYTRVNLAEMGRAKGELOVHCNEVLHVT	706	Db	824	QERDAARARQLAGAEQ-----EGKTALEOQRAAHKEVNLREWEKERSWHOQBLAKE	878
Db	2927	QAMQKGLIVREHGTRILLEAQATGGYIDPVSHRVFPUDVAYR--RGYFDEEMNRYLADP	2983	Qy	617	MVDYEASPLEPKAVLEDTTLEAVGLLRRVD-----GCCLLSVVKVNTDGYKRL-LQDL	668
Oy	707	DTMFQGGCWNHARYNVSYTMKDTAAHTGTTIPYNSRAQQQLIALID-MTOQCC--TVTRKPS	763	Db	879	SLERERKE-LEMRLKECQTEMEMAQACREERETQAEESALCQMCOLETEKERYSLLETLTQ	937
Db	2984	SDDTKGEFPDNTHENLYL-----OLLERCVEDPETGLCLLPLDKA	3026	Qy	669	EAKVATSGDSFY-----IRVNLAEMGRAKGELOVHCNEVLHVTDTMQGCCGWHAIRVNS	723
Qy	764	SGCPQKLIVRIVSMKDKAKASPLRLSFDRGOLDP--SRMEGSSTCFCWAESCLTLPVTLVWP	821	Db	938	QKELADASQSLERLQRQDMQVQLKEQFTTGTGILQTOQCE-----AQR-----	978
Db	3027	RKG-----ELYTD---SEARDVFVKATVSAFPGKFQGKVTIW-EIINSEYFT-----	3071	Qy	724	YTMKDTAAHGTIPNYSRAQQQLIALID-MTQOCCTV-----TRKPSSGGPQLKRVTSMDK	778
Oy	822	HRPARPPVLPVPRAVGKILSEKLLCQGFKKCLAYLSQEEYAMSQRGDIQEG	877	Db	979	-ELKE-PARQHDDLLAQEESSSLQDQMLQKQVDEDLKSLVQADSDQRLEQEYQEKE	1036
Db	3072	--AEQRDRDLRQFRGRITVEKII-----KIIITVVVEEE-----QKSGRLCFEG	3113	Qy	779	AKASPLRLSFDRGQLOPDSRMESGSTCWAEASCLTLYPTYLWPHRPARPPVLPVRAVG	838
Db	1037	LRETO--EYNRIQEKELEREKASLT-----LSLM-----EKEORLVLQI	1075	Qy	839	KILSEKUCLL-----QGFKKCLL-----EYEAWSQRQDIIQEGEVSGSRC	884
Qy	7	T08621	centrosome associated protein CEP250 - human	Db	1076	SIROQELSALRQDMQEAQEQKELSAQMELLRQEYKEKEADFLAQEQLLLEAS	1131
C.Species:	Homo sapiens (man)		C.Species: Homo sapiens (man)	C.Species: Homo sapiens (man)			
C.Date:	11-Jun-1999	#sequence_revision 11-Jun-1999	#text_change 21-Jul-2000	Qy	885	WVTRHAYESLMBEKNTHALLDYOLDSVCTLHMDIFPIVHVSNERMAKKLKGLORLGT	944
R.Mack,	G.J.:	Rees, J.; Sandblom, O.; Balczon, R.; Fritzler, M.J.; Rattner, J.B.	Arthritis Rheum. 41: 551-558, 1998	Db	1132	-----HTEQQLRQLRSLWAQEAQANQLH-----LRLRS	1158
A.Title:	Autoantibodies to a group of centrosomal proteins in human autoimmune sera react with centrosome associated protein CEP250 - human		A.Title: Autoantibodies to a group of centrosomal proteins in human autoimmune sera react with centrosome associated protein CEP250 - human	Qy	945	SEEQL-LEAARQEGIDLDRAPC----LYSSL-----APDGWS---	976
A.Reference number:	216462;	MUID: 98165428	A.Reference number: 216462;	Db	1159	TBSQLELAAECPGNAQAOQASLQASLGVSCESRPELSSGGDSAPSVMGLEP	1218
A.Accession:	T08621		A.Accession: T08621	Qy	977	DLDGLLSCAVR-----QIADECKKV-----WVTEQSPr	1004
A.Molecule type:	mRNA		A.Molecule type: mRNA	Db	1219	DONGARSLFKRGPPLTALSAAVASNLKHODLWKTTQQT	1259
A.Residues:	1-2442 <NAC>		A.Experimentally determined source: cell line HeLa	RESULT	8		
A.Cross-references:	EMBL:AF022655; NID:92832236; PIDN: AAC06349.1; PID:92832237		A.Experimentally determined source: cell line HeLa	I54378			
Best Local Similarity	4.8%	Score 248.5; DB 2;	Length 2442;				
Matches 221; Conservation	22.1%	Pred. No. 5.5e-05;	Indels 299; Gaps 47;				
Mismatches	140; Misnatches	34;					
Indels	299;						
Qy	123	KLT--ECLAGAIGSLSQELNEQKOEVKLLRRCQOLQHEGLAETRAEGLH-----QL	173				
Db	439	KLTGERDTLQAGTQVTDQGEVLSKRELIQKAREELRQOLEVLEQAWRLRNVNVELQ	498				
Oy	174	EADHSR-MKREVSAAHHR---EVLRKDEMMSLHSYNSAJOKEEIAASRSRLOEELY	228				
Db	499	QGDSAQGQKEEQQELHLAVFERERLQOEMMLEAKQSESISELITLREALIESIHLEGEL	558				

A	Molecule type:	mRNA
A	Residues:	1-1116 <RBS>
A	Cross references:	GB:L27476; NID:9498012; PIDN:AAA61300_1; PID:9498013
C	Genetics:	
A	Gene:	X104
C	Superfamily:	guanylate kinase homology; GLGF domain homology
F	1-116/Domain:	guanylate kinase homology <GU1>
F	737-879/Domain:	guanylate kinase homology <GK1>
Query Match	4.8%	Score 246 5; DB 2; Length 1116;
Best Local Similarity	21.2%	Pred. No. 2.5e-05;
Matches	212;	Conservative 137; Mismatches 393; Indels 257; Gaps 1
Qy	55 EEVYHLSPRLTNSAMRAGHLLDLKLTRKGNGAAIAFLESLKHPNP -	DVYTLYTGLOPDVDF 1
Db	94 EIVLHS -FAVQQLRKSGKVAIAVVKPRKVQAALQA ---	SPPLDODDRAFEYNEDEFDG 14
Qy	113 SNF -SGLMETSKLTTECLAGAIGSQQEELNEQKQKEVLLERRCQOLOEHUGLAETRAEGLH 1	
Db	149 RFSRFSYSERLRSN HGGGRSRSSW -SPERGRAPH -	- ERARSNERDLSRDRSRRSLS 20
Qy	172 Q -LEADHSRMKREVAHSFHETYLRLADEMISL -HYSNALQEELAASCRSLQSOLEYLL 21	
Db	203 RGLDQHQHARTD-----	RSRGRSLERGDHDFGPERRDQRDRSRRSRSID----- 21
Qy	230 KQELQRANMYSSCCEBLEQEOSLRTASDQEESDEELNRLKEENKLRSLTFLSAEKRDILEQ 22	
Db	247 -QDYERAY -----	HRYADPYER-----AY 22
Qy	290 SLDGEAGRSRQFLVERHSLDERAYAERAEQWEEKEOTLLQPKSKAACQCYREKINA 30	
Db	266 SPEYRGAR -----	HDAASRG -PGRSRSEHPPHSRSRSPSPERGRGPPIGVLLMKMSRAN- 31
Qy	350 LQAQVCELQKERRQAYSARDSAQ -- -RETSQS -LVEKDSLRQ - -VFELTDQVCE - -LR 40	
Db	317 -----	EYIGLRLGSQIVKEMTRGQATKDNLHEDTIILKINGVTENNMSL 31
Qy	401 TQLRLOQAEPPGVLKQEARTRPCEPREKQRLVRMHAICPRDD - -SDCSLVSSTES - - - 45	
Db	364 TDARKLIEKSRLK -QLVYED -----	SQTLNIPSLNDSNDSDEIDESTSFSEPSPEE 41
Qy	454 - -QJLSDLSATSRSRFLDSFRSSPSSAPPSSQSLXKRVADDFGPPWSSCLEPEPGDQ 51	
Db	419 RRHQISDYDHSSSEKLKERPSRSRDTPSRLS -----	RMGATPTPKS -----TG 46
Qy	512 ALPGAKAGDBHLDEYELLDPLDQPLESSLOPVSPGRLDVSSESVLMMRRPARIL - -SQV 56	
Db	464 DIACTGVPETKNCPEYQEPPEAPP -----QPKAPR -----TFLRSPDENDYGNNT 51	
Qy	570 TMLAF -QGDALLEQTSVIGLNLTGFIHRTVPGAADOMALPCTGTYDPEASEPLFK 62	
Db	511 KMVRKKGDSV - -GURLAGNDVIFIVAGIQEGTSAAEQGLQGDQILKVNTODFRGL - -	56
Qy	629 AVELEDTLEAVGLLRRVDG - FCCLSVKVNTDGYKRLQDLEAKVATSGDSFYIRVNLM 68	
Db	567 -VREDAVLY -- -LLEIPKGEMTILASRSDAYRDIL -----ACGRGDSFTRSFEC 61	
Qy	688 EGRAKGELOQYHCNEVHLYHTDMFOG - CGCWHAHRVNSYTMKDAAHTGTPNYSRAQOOLI 74	
Db	616 EKEPQSLAFTGEVFRVYDTLGKLGNLWAVRIGNELEK -----GLIPNAAEQ - -M 66	
Qy	747 ALIQDMTQOCCTVTRKPKSSGGPKQKLIVRYSMDKAKASPLRLSPFRGOLDPSMEGSSTOFW 80	
Db	669 ASYON -----AGDRNAGDRAFWRM ---	-RQRSGWYKKNLRSREDLTAVSVSTKFP 71
Qy	807 AESCLTLYPDTWLPHRPARPVYLLPRAVGKILSEKLCLLQGFKKCLAETLSQSEBEYA 86	
Db	718 A -----YERYLRLREAGEKFKRVVLF - GPIADIAAMEKLA - - -	- - - - -NELPD 75
Qy	867 WSQRGDI TQEGETSGGWRWTR HAVESLMEKNTHALDVQLDSVCTLHRMDIFPVLTHV 92	
Db	754 WEFQTAKEPKDAGSEKSTGQVVRNLYTQVQIEQDKHALLDVTPKAVDLNYTQWSFVISF 81	

Qy	926	SVNEKMK-----ARKLKGQLRGTSFQOLLEAARQEGLDLDR	962
Db	814	TPDSRQGVNTMRQLDPTSNNSRRKLFDANKLKKTCALIFTATINL-----	860
Qy	963	APCLYSIPLADPQWSDLGILSCVROTADEQQKVWTEQ 1001	
Db	861	-----NSANDSW-----FGSLKDTHQGEAVWVSE 887	
RESULT 9			
T4737	myosin II heavy chain [imported] - Naegleria fowleri (fragment)		
C;Species:	Naegleria fowleri		
C;Date:	20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000		
C;Accession:	T47237		
R;Shaw, D.R.; Sullivan, P.K.; Mariano-Cabral, F.; Ennis, H.L.			
submitted to the EMBL Data Library, December 1995			
A;Description: Codon usage in Naegleria fowleri.			
A;Reference number: Z24413			
A;Accession: T47237			
A;Status: preliminary; translated from GB/EMBL/DBDJ			
A;Molecule type: mRNA			
A;Residues: 1-746 <SHA>			
A;Cross-references: EMBL:U43192; PIDN:AAB01786.1			
A;Experimental source: strain LEE mp; cell type amoeba			
Query Match	4.8%	Score 245; DB 2; Length 746;	
Best Local Similarity	19.2%	Pred. No. 1.7e-05;	
Matches 160; Conservatism 151; Mismatches 293; Indels 228; Gaps			
Qy	47	KVLCQLDDEFEVLSPRITNSAMRAGHLLDKLRKGNGIAFLESLSKHFHNPDVYTLVTLG	106
Db	19	KKLSESSKDELTLQLNKTND-EKNEYLVNLKKAED-----LKNLKSDD-----L	64
Qy	107	QPUDFSN----FSGFMETSKTLCIAGAISLQOEENIQEGQK--EVLL----	150
Db	65	QAEKDDDSNRIRKLEODREKEPOLSENLAKRIDLNEARTKEAQKSTEMELSSVKDDL	124
Qy	151	--RRCQOLQEHGLAETRAEGLHLEADHSRMKREVEAHFHEVRLRKDEMLSLSLYHS	206
Db	125	NPTKQRQAFQQLSDLEAERANLENLSDTEGGKQDQSDFQQL-----Q	170
Qy	207	NALQEKLAAASCRSRLSQEELYLKKQELQRANHYSSCCEFLQEOQLRTASDOESGDEEL-N	265
Db	171	NEQNTERINLQNMKSEERL-----OQELEMK-RSLQDNKONESLSDS	213
Qy	266	RLKEENKEKRSLETFSLAEKDILEQSLDEARGSTOELLVERIHL-----RERAVAEERQRY	322
Db	214	KVSKSLEKIRELTALLETERSSTTDLKQSKMDKEYKRLAQQLQETEQALIGETOKK--	271
Qy	323	WEEKQTLLQFOQSKMKAQOLYERKVN-----ALQAQYCEL	357
Db	272	-NDADRNKYQKOLESELQVKSERDRLNKDLNNTSGDMANLKRQLDENSNLVAKKAEQKL	330
Qy	358	QKERDQAYSARDSAQREISQSLVERYFELDTQVCELRTOFRQ-----LQAEP	410
Db	331	QKDLSDHGDREETEEQI-----DLRKQJQELTSRLSNDANOTKQSRQNLSEEN	383
Qy	411	PGVLKQFARTTRECPKPREKQLVR-MHAICPRDSDCSUVSSTESQSLSDLSATSSRLESD	469
Db	384	NRLKSEVSRLREDLQNRRLUQEMEVQSESENE-----KSELJTLQ-----	427
Qy	470	SFRSSSPAPPSSQSLSYKRNAEDFGEEFWSFSSCLETEPGDPALPGKAGDPHDYELLD	529
Db	428	-----KLOEADEVKSEVFKELKD---LSKNASSGGGGVVGVDFSEVEVKURREPN-	471
Qy	530	TADLPOLESSQQLQPSVPSRQDVS-----EGVLMRMRPARRISQTMIAFGQ	576
Db	472	--QLAQLKARVEEVTKQRDVENKKRSVEMDLTEMKTRLQTEERLRKVKVEQQKSYMEC	529
Qy	577	DALLEQSVIGGNLTTGIFHRVTPGSAADOMALRPSTIVMDYFA-SEPLFKAVLFD--	633

Query	Match	Score	Length	DB
530 DEIRE-----	-----LAEEAEDLDELNRKLEHQALIQOLRQLQERH	569		
634 -TITLEAVGLRRVVDGFCCLSKVNTGYKRLLQDKERKATSGDSFYIRVNLAAMEGRA	691			
570 SRAAEESATRQKR-----EIEELQDLEERA-----KLDEA	603			
692 KGEIQVHNEVHLVHTDTMFGCGCWHHRVNSTMKD-TAAHGTIPNYSRAQQQIALLIQ	750			
604 RRLKQYQEINELDNQTAQ-----AKERSAASRDMKKADRLREIORRQEEARAKQ	657			
751 DMTOQCTVTRKPSSGGPKPLVRYISMDKAKSPLRLSFDRGOLDPSRMEGSS	802			
658 DLQRQLTKVEREN-----KLLQ-----SQSODAS-----KYOKAEDQEKRORLAEN	698			
RESULT 10				
T30171				
ninein - mouse				
C;species: Mus musculus (house mouse)				
C;revision: 22-Oct-1999 #sequence_revision 22-Oct-1999				
C;Accession: T30171				
C;Accession: T30171				
M.; Moudjou, M.; Ferguson, D.J.P.; Mucklow, M.; Belkaid, Y.; Milon, J. J. Cell Sci 109, 179-190, 1996				
A;Title: Molecular characterisation of ninein, a new coiled-coil protein of the centrosome				
A;Reference number: Z20751; MUID: 96431720				
A;Status: preliminary; translated from GB/EMBL/DDJB				
A;Residues: 1-2168 <B00U>				
Cross-references: EMBL:U04342; NID:91113864; PID:AAA83234.1				
A;Experimental source: strain C57BL/6				
A;Note: localised specifically in the pericentriolar matrix of the centrosome				
Query Match 4.7% Score 243.5; DB 2; Length 2168; Best Local Similarity 19.6%; Pred. No. 8.6e-05; Matches 246; Conservative 183; Mismatches 380; Indels 443; Gaps 61;				
QY 13 ALDPETLWMMESHRHRVRCIPSRLTPYLQRQKVLQCOLDEEVLHSPLTLNSMRAGH	72			
QY 639 SIEELVIEQMKFQHH---RDLCIHLR-----ELEDYKVRHYEQQLDPRVASEQ	84			
QY 73 LIDILIKTRQKGIAFLESLKFHNPDIVYTL---VTGLOPDV-DFSNEFSGLM-----	119			
QY 685 EQAMKQKVEQ-----VHLEKRVSELRSEIADLEGQAAVLRREAHHKASC	730			
QY 120 ---ETSKLITECLAGAIGSQQEIQNQEKQEVYLRRQCOL----QEHIGLA----ET	165			
Db 731 RHEEKEKROLQMAFDEEKAQQLQELRQE--HEREIQLRQQAETFRQEREGLAQAAWTEE	788			
QY 166 RABELHQ-----LEADHSMRREVSAHFEVYLRLKDEMMSLHSNALQE KELAA	216			
Db 789 KVRSLEQSYEQQLISLEEKHALKEE-----LREB--LSEHHRRRLQSCREEME	835			
QY 217 SRCRSLQEEPYLHQELQRANMYSSES--LELOQOSLRTASDQESGDEELNRLKEENEKL	274			
Db 836 TECNRVRSQI-----BAQCDQDCYEKTHCQTLQSLEVRHR--QELRLLDQHLE	885			
QY 275 RS-LTFSLAEKDILEQSLDEARGSRQELVERIHSRLRERVAAREQEQ-----YWE	324			
Db 886 RSQWEF--EKDELAQEQCTDAQEQIKEALQ----RERATAAMKQEQELTERYKDRLN	937			
QY 325 ---EKEOTLQFQSKMCAQQLYKREKVNQALQAYCQELQERDQAYARDSAQ-----	372			
Db 938 ILSTEREQLLQDIDLQNAHSE--SQHGLLSQGLEL--KRSQERELRQGQLCQTGVS	992			
QY 373 -----REISSLVKEDSLRRQVFELTDQ-----VCELTQRLQ	405			
Db 993 EQLAPSEPELRLRYHEQERREMGTKLALESAHASLASLERAQEAEMSTEICRLONTVKD	1052			
QY 406 LQ-----APSPGVL-----KQEA 418				
Query Match 4.7% Score 242.5; DB 2; Length 3225; Best Local Similarity 16.3%; Pred. No. 0.001; Matches 220; Conservative 194; Mismatches 396; Indels 331; Gaps 41				

Qy	9	SALTALDEE--TLWEMMESHRRHRIVRICPPSLRQLVAKVLCPOLD-EEFVLLHSPLRLT 64	Qy	954	RQEGLDLDRAPLSSL-----APDGWSIDLGLLSCVRQIADEQKVVVTE 1000
Db	703	SAFTALSERDOLLSQYKELSVATELRAQV- ⁻ QLEMLMSEAERQRRLDYESTAHDNLLT 761	Db	1606	ROPKQE----LYKURIASTEANKKEPEKQLOAEQMEEMKEKMRFAKSQQKILELE 1659
Qy	65	NSAMRAGHLDDLLRKTRGNGAAFLS-----LKHFNPDVYTLYTGQLQPDYDFSNFSGL 118	Qy	1001	Q 1001
Db	762	EQT-----HSLS-LEAKSKDVKEVLQNEDLVQLOFSEQS-TLIRLSQSQI-----807	Db	1660	E 1660
Qy	119	METSKLQEELNECLAGA-----IGSQEELNQEKGKEVLLRRCOOLQENGLAETR-AEGHL 171	Qy	12	RESULT 12
Db	808	--NKESEVLEGAERVHRHISSEVEILSQUALSOKDELETIKMDL-----LLERKRDYETLQ 859	Db	A56539	AS56539
Qy	172	QLEADHSRMKREVSAYAHF-EVLRKDEMMLSLSHYSNALQEKEALAASCRCSLQEELYLL- 229	Qy	1329	giantin - human
Db	860	QTEERKDQVTTELSFSMTEKMYVQLNEEKFSLGV-----EIKTLEQNLNLSS 905	Qy	1329 < SEQ>	N: Alternative names: macrogolgin
Qy	230	---KOELQRANMVSS-CETELQEBOSLRTAQESGDEEUNRLKEENEKRSLSLTSFLSA 282	Db	19-Oct-1995	C: Species: Homo sapiens (man)
Db	906	RAEEAKKEQVEDDENEVSSGLKQNYDENSPAGQISKEELQHEFDLKKENEKQRK-----958	Qy	1329	C: Accession: A56539; MUID: 94187728
Qy	283	ERDILEOSLERDERSRQLPVLTHSRERAV-----AAEROEQWEEKKOTLL 331	Qy	1329	A: Molecule type: mRNA
Db	959	--RKLQALINLNRELLQVNRQNSRLEELANLKDSEKKEIPLSETERGEYEDEKENKEY 1012	Db	1329	A: Residues: ENBL:X75304; NID: 9405714; PIDN:CAA53052.1; PID: 9405715
Qy	332	QFOKSKMKAQ---LY-REKVNAHQAOYCVELQKERRQDAYSARDAQ---REYQSILVEKDS 384	Qy	1329	A: Cross references: EMBL: X75304; NID: 9405714; PIDN:CAA53052.1; PID: 9405715
Qy	1013	SECVCTSKCQEETYLKQTISKEVELQHIRKDLERKLAEEQFQALYKVNNTLQDK-- 1070	Qy	1329	C: Genetics:
Qy	385	LRRQVFELTDQCELTQTLRQLQA-----EPPGVLKQEARPREPCPREKQ 429	Db	1329	A: Gene: GCP371
Db	1071	--TNQDILQAEITSENQAIQKLITISNTDASGDGSVALVKETVVISPPCTGSSE 1122	Db	1329	A: Cross references: GDB: 451958
Qy	430	RIVRMHAICPDRDCSLSVSTESLSDL-TATSSHEL-----DSFRSSPAPPSQOS 483	Qy	1329	A: Map position: 3913-31-3Q13.31
Db	1123	HWK-----PELEFKILALEKEKEQOLQKQLEALTSRKAILKQAKERHREELKQKD 1176	Db	1329	C: Superfamily: giantin
Qy	484	LYKRAEADFGEFPWFSSC-----LEIPEDPGPALGAK---AGDPHDYELLDT 530	Qy	1329	C: Keywords: coiled coil; Golgi apparatus; transmembrane protein
Db	1177	DNRLLQDFDSEKSKENNINQDOLRQLQIVQRESIDTIGLPSTQEQCSSTSPLFKA 1236	F: 3238-3254/Domain: transmembrane #status predicted <TMN>		
Qy	531	AD----LPQLESSLQPYSP-----GRDYESGVLMRRPARI-----LS 567	Query	9	SALTALDEE--TLWEMMESHRRHRIVRICPSRLTPYLRQAKVLCQOLD-EEEVLHSPRLT 64
Db	1237	TQHQHTHPVLESNLCPDPWPSHEDASLOGGTSVAQIAKQIREIEAKVEELKVSSTS 1296	Db	737	SAFTALSERDOLLSQYKELSVATELRAQV-KOLEMNLAEARORRLLYESQTAHDNLLT 795
Qy	568	QVMTMIAFGDALLQFISVIGNULGTIFHRVTPGSAQDOMARPTQIVVMDYEASEPLF 627	Qy	65	NSAMRACHLLDLIKTRKGNGATAFLS-----LKFHNPDVYTLVTCQDQDFDSNFSGL 118
Db	1297	ELTKKSEEVFOQEQI-----NKQGLEIESL-----KTVSHAEVHAESL 1336	Db	796	EQJ-----HSLS-TEAKSKDVKEVLQNLDDYQLOFSEQS-TLITRSLSQI-----841
Qy	628	KAVLDEDTLEEVAVGL-LRRVGFCCUSLKVNTGYKRLLDLEAKVATGSDSFYIRVNL 685	Qy	119	METSKLTKECLAGA-----IGSQEELNQEKQOREVLLRQCOOLQELGAEIR--AEGLH 171
Db	1337	QKLELQSKLQI-----ELQKLDLDEQKLISKKEV-----YLSQL 1381	Db	842	-----NKESEVLEGAERYVRHISQKVEELSQALKSOKELETKMDSL-----ILLEKKRDYETLQ 893
Qy	686	AMEGRAKGELOVHCNEVLYHVTDMFOCGCWMAHRYNSYTMKDTAAHTGIPNSRAQQQL 745	Qy	172	QLEADHSRMKREVSAYAHF-VLRLKDEMMLSLSHYSNALQEKEALAASCRCSLQEELYLL- 229
Db	1382	SEKEAALTKQ--TEIEQEDLI-----KALHTQLEMKAEBHD 1419	Db	894	-----QTEERDQQVTEISFSSTERVQLNERFKFLSV-----EIKTLEQINLNS 939
Qy	746	IALIQDMTQOCUTVTRPKSSGGPKQLVRLIVSMKAQASPLRISFDRQDLSMEGSSTCF 805	Qy	230	-----KQELQRANMYS-CELEQEOSSLRTASDQESGEELNRLKEENEKLRSLTFTSLA 282
Db	1420	IQLQ-----VELCMLKOPPEEIGEESRAKQQLQKQALI-----SRKEALKE-- 1464	Db	993	-----RKQLQALINNKELLQVRSLLEELANKDESKKEIPLSETERGEVEEDKENKEY 1046
Qy	806	WAESCLTLPVPTLWPHPRPARPVLLVPRVAVGKILSEKLLQGPFRKCLAEYLQSBEEYE 865	Qy	332	QFQKSXKMACQ---LY-REKVNAHQAOYCVELKDERQDAYSARDSAQ--RETSQSLVYKDS 384
Db	1465	-----NKSQOEELSLAARGTIERBLTKSLADVESQ 1492	Db	1047	-----SERKVTSKQCQEPIYKQTTISKEVELQHRLKDEELAAEEOFQALVKMNQTLQDK-- 1104
Qy	866	AWSORGDIQEGEVSGGRCWYTRHAYESLMEKNTHALLDYQ-LDSVCTLHMRDIFIVIH 924	Qy	385	-----EPGPVLKQEAUTREPCPREKQ 429
Db	1493	VSAQN----KEQDTVGRALLQEEQRDKLITEMDRSLLENSSCSLKLALEGL-- 1545	Db	1105	-----TNQDILQAEISQNLQITISNTDASGDSVALVYKETWVSPPTGSE 1156
Qy	925	VSVNEKMAKKL-----KGLQR----LGTSF-----EQLLEAA 953	Qy	430	-----DSFRSSPAPPSSQS 483
Db	1546	TEDKEKLVKETEISLKSXKIAESTENQEKHKELQKEYEILLOSYENVSNEARIQHVVEAV 1605	Db	1157	-----PELEEKTLAELKEKEQOLQKLOEALTSRKAILKKQAEKERHUREELKQKD 1210

Qy	484 LYKRAEDFGEEPWSSC-----LEIFEGDPGALPGAK-----AGDPHDYELLDT 530	Qy	484 LYKRAEDDFGEERPWSFSCLEIPGDDGALPGAKAGDPHDYELLDTADLPOLESSLQPV 543
Db	1211 DYNRLRQEFDQSKEENIGDQLRQLQIVQEVESIDQEQCSLPLSPGLEPLFK 1270	Db	418 -----RMGAMPTPFSTGDI-----ATRAVTVVDAEKPKYQDDIVAVOPKAVRTI 464
Qy	531 AD---LPQLESSLQPVSP-----GRDVSSEGVLMRRPARRI-----LS 567	Qy	544 -SPGRLLDVSESGVLMRRPARRILSQQVMTLAF-QGDALLEQTSVIGGNLTGIFHRYTPG 601
Db	1271 TEQHTTQPVLSENLCPDWPSHEDASALQGTSVAQIAKIEAEKVELKYSSTS 1330	Db	465 LKPSPEDEAIG-----PNTKMRERKGDSV--GLRLAGGNDVGFIAGIQEG 510
Qy	568 OVTMLAFQGDALLEQISVIGGNLTGIFHRYTPGSADQMLARPGHQIYMDYEASEPLF 627	Qy	602 SAQDMQMLRPQGTOIVMDYEASEPLFAVLDFTLBEAVGLLRRV--DGFCCLSYKVNT 658
Db	1331 ELTKKSEEVPLQEQI-----NKQGLEIESL-----KTVHEAEVYHAESL 1370	Db	511 TSADQBELQEGDQQLKNAQD---FRGIVR---EAVDLYLEIPKGD----TVTILA 557
Qy	628 KAYLDEEFTLEBAVGL---LRRVGFCCLSVSKVNTDGYKRLQDLEAKVATGSDSFYIRVNL 685	Qy	659 DGYKRLQDLEAKVATSGSDSFYIRVNNAMEGRAKGEQVHCNEEVLVTDTMFQG-CCCWH 717
Db	1371 QOKLSSQLOQA-GLEHLR-----ELQPKLDELQKLSSKEEDVS-----LGSQ 1415	Db	558 QSKYEYRDINA--CGRGDSFFTRSHCECERESPQSLAFTREIFRVDTLYDGKLGWL 615
Qy	686 AMEGRAKGKGELQVHCNEEVLYHTDTMFQGCGWHAHRVNSYTMKDTAAHGTIPNYSRAQQQL 745	Qy	718 AHRVNSYTMKDTAAHGTIPNYSRAQQQLIQLQDMTCQCTVTRKPSSGGPKQLVRIVSMD 777
Db	1416 SEKEAALTKIQ---TEIIEQDIL-----KALHTOLEMKAEDHER 1453	Db	616 AVRIGNELEK----GLIPNRSAEQ--MASVONAQKD-----GPS-----D 650
Qy	746 IALIQDMTQQQTVTVKPSSGGPKQLVTRIVSMKDAKASPLRUSFDQQLDSRMSSESSTCF 805	Qy	778 KA-----KASPLRUSFDQQLDPSMEGSSCTCFNWCESLTLPVTPYLWPHRPARRPV 830
Db	1454 IKOLQ---VELEMCKQKPEEIGEESRAKKQIQRKQALI-----SRKEALKE-- 1498	Db	651 RADFWRTRGQRSGVKNNLRKSREDLTAVSVGTKFPA-----YERVOQIREAGFRPRV 702
Qy	806 WAESCLTLVPLVWPHRPARRPVPLVWPHRPARRPVPLVWPHRPARRPVPLVWPHRPARRPV 865	Qy	831 LLVPRAVGKILSEKLC----LQGFFKCLAEYLSQLSEYEAMSQRGDIIQGEVSGRCW 885
Db	1499 -----NKSLOEELSALARTRTIERLTKSLADESQ 1526	Db	703 VTF-GPADVAMEKLSLDLPHLYQTAK-----IPERDAGSEKATGV----- 742
Qy	866 AWSQRGDIIQGEVGGRCRWVTRHAYESLMEKNTHALLDVQ-LDSVCTLHMRDIFPIVH 924	Qy	886 VTRHAYESLMEKNTHALLDVQDLSVCTLHMRDIFPIVTHVSNEVKMKLKKLGLQRL--- 942
Db	1527 VSAQN---KEDTIVLRLALLQEERDKLITEMDRSLLENQSLSSCESLKLAL EGL--- 1579	Db	743 VRLNTYFQIIEQDKHAILDVTPKAVDLNLYTQWFPIVFFNFDSKQGVKTMR--QRLCST 800
Qy	925 VSVNEKMAKKK-----EQLSE-----KGLQR-----LGTS 953	Qy	943 -GTSEEDLLEEARQEEDGLDR--APCLYSSPLADPGNRLDGLLSCVRAIADEQKVKWWT 999
Db	1580 TDEKELVKEYESLKSKKIAESTENQEKHLQKEYEILLQSQYNSNEARIQHVVEAV 1639	Db	801 SNKSSRKYEQANKLKLRTCSHLFTATINLNNSANDSW-----YGSLRKTIOQQGEAVWW 854
Qy	954 RQEGLDRLPACYLSS-----APDGWSLDGILSCVRSYRQIADEQKVKWWT 1000	Qy	1000 EQ 1001
Db	1640 RQEKE-----LYGKLURSTBANKKETEKQLEAQEEMEENKEKMRKFAKSQQKILEE 1693	Db	855 SE 856
Qy	1001 Q 1001	RESULT 14	
Db	1694 E 1694	I46236	tight junction protein - dog (fragment)
		C;Species: Canis lupus familiaris (dog)	
		C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 16-Dec-1998	
		C;Accession: I46236	
		R;Jesaitis, L.A.;Goodenough, D.A.	
		J. Cell. Biol. 124: 949-961, 1994.	
		A;Title: Molecular characterization and tissue distribution of ZO-2, a tight junction	
		A;Reference: A54475; MUID:94179414	
		A;Accession: I46236	
		A;Status: preliminary; translated from GB/EMBL/DDBJ	
		A;Molecule type: mRNA	
		A;Residues: 1-775 <JES>	
		A;Cross-references: GB:L27152; NID:9463047; PID:9507892	
		C;Gene: ZO-2	
		C;Superfamily: guanylate kinase homology; GLGF domain homology	
		F;316-464/Domain: guanylate kinase homology <GLG3>	
RESULT 13			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.7%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 22.2%	Score 242; DB 2; Length 1163;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000	Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		
C;Accession: JE0366			
R;Collins, J.R.; Rizzolo, L.J.			
Biochem. Biophys. Res. Commun. 252: 617-622, 1998			
A;Title: Protein-binding domains of the tight junction protein, ZO-2, are highly conserv			
A;Reference number: JE0366			
A;Accession: JE0366			
A;Status: Preliminary			
A;Molecule type: mRNA			
A;Residues: 1-1163 <QOL>			
A;Cross-references: GB:AF085184; NID:93820579; PID: AAC95469.1; PID: g3820580			
Query Match 4.7%			
Best Local Similarity 22.2%			
Matches 147; Conservative 102; MisMatches 239; Indels 174; Gaps 30;			
JE0366	tight junction protein, ZO-2 - chicken	Query Match 4.6%	
C;Species: Gallus gallus (chicken)	Best Local Similarity 23.5%	Score 236.5; DB 2; Length 775;	
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 16-Dec-1998	Matches 133; Conservative 86; MisMatches 231; Indels 115; Gaps 26;		

Dp	54	VGTTENSKEPRYQEE-----PPAPQPKAAPR-----TFLRSPSEDEAIFYGPNTK	96	C;Keywords: actin binding; ATP; coiled coil; hydroxlate; methylated amino acid; nucleo F;84-764/Region: myosin motor domain homology <MMOT> F;174-181/Region: nucleotide-binding motif A (P-loop)
Qy	571	MIAF-OCBALLEPISVIGGNLIGIFIFTRVTPSAADOMALRIGVTFMVDYEASEPFLFKA	629	F;552-616/Region: actin binding #status predicted F;626-640/Region: actin binding #status predicted F;837-1936/Region: coiled coil #status predicted <COI>
Db	97	MFRFKGDSV--GLRLAGNDVGFVAGIQESTSAAEGLQGDLKVNQDFRGL--	151	F;837-1277/Region: S2
Qy	630	VLEDTTLEAVGLLRRYDG-FCCLSVKVNTDGYKRLLQDLEAKVATSGDSFYIRVNNAME	688	F;11937-11959/Region: light meromyosin
Db	152	VREDAVLY----LLEIPKGEMMTLAQSRADYRDIL-----ACGRGDSFFIRSFECE	201	F;125/Modified site: N6, N6'-trimethyllysine (Lys) #status predicted
Qy	689	GRAKGEQVHNEVLYHTDMFOG-CCEWHARVNVSNTMKDTAAHGTIPNYSRAQQOLIA	747	F;180/Binding site: ATP (Lys) #status predicted
Db	202	KETPQSLAFSRSQEVRVYDFTLIDGKLGHWLARIGNELE----GLIPNKSRAEQ--MA	254	F;694-704/Active site: Cys #status predicted
Qy	748	LKDMTOCCTVTRKPSSCGPQKLVRITYSMDKAKASPIURLSFRFGDPSMEGSSTCFWA	807	Query Match 4.6%; Score 235; DB 1; Length 1959;
Db	255	SQN-----AORDNGDRADEFWRM---RQRSGMKNKLRSREDLTAAVSVKFPA	303	Best Local Similarity 20.8%; Pred. No. 0_00021; Matches 244; Conservative 170; Mismatches 385; Indels 376; Gaps 58;
Qy	808	ESCLTVLPYTLYNPHRPARPVYL----VPRAVGKILSEKUCUCLLGQFKKCLAEYLSQEE	863	Qy 14 LDEETLWEMMESHRH--RIVRCICPS----RLTPYLRQAKVLQDDEEVYLHSPPR-- 62
Db	304	-----YERVILREAGFKRPVVLFGSPIADALEKIANELDLFQTAK----TEP	347	Db 649 LYKEQLAKLMLATURNTNFVRCIIPNHERRKAGKLDPH---VLDQIICNCVLEGTRIC 704
Qy	864	YEAWNSQRDIIQEVESSGRCWYTRHAEVSLMEKNTHALDVQDSYCTLHMFDIPIVI	923	Qy 63 -----LTNSAMRAGHLLDLKTRKGNGKNAIFLESLKPHNPDVYTL- 102
Db	348	KDAGSERSSGV-----VRLNTVQRIQLEQDKHALDVTPKAVDNLNTQWFPPIV	396	Db 705 ROGFPNRYYVFQEERFORYEILTNAAIPKGFM----DGQACVLMKALEL-DSNLRYIG 757
Qy	924	HVSNEVKAKKLRKGLORLGTSEQLLEAORREGDIDRAPDR-----SSLAPGWS	976	Qy 103 -----VTGQDQDVDFSFHSGLMLMETSKLTECLAG-----AIGSLQEELN 140
Db	397	FNPDSROGVKMR-QRLNPFSNKSSRKLYDOANKLKKT-CAHLFTATINLNSANDSW-	452	Db 758 QSKVFFERAVLHLEEBFDL-----KTDVILQFACCGRVYLARKAFAKRQQLT 807
Qy	977	DLGGLLSCVROQIADEOKKVWVTEQ 1001		Qy 141 QEKGQKEVLLRRCQLOEHGLAE-----TRAEGLHQADHSRMKREVAHFHEVLR 193
Db	453	-----FGSLKDPTQHQOGEAWVWSE 472		Db 808 AMK---VLRQNC--AYKLKRNWQWWRFLFKVKPLLQV----SROEENMAKEEELIK 856
Qy	194	LKDEMLSLAYSNALQEQKELAASRCSRSLQEEELYLKQELQRANMYSSCELQESLR-	252	Qy 194 LKDEMLSLAYSNALQEQKELAASRCSRSLQEEELYLKQELQRANMYSSCELQESLR- 252
Db	857	VIERQLAA---ENRLSIME-----TFAQQLMAEKMOLQE-QLQAELCAEAEERA	904	Db 857 VIERQLAA---ENRLSIME-----TFAQQLMAEKMOLQE-QLQAELCAEAEERA 904
Qy	253	--TASDQESGDFEL----NRLKEENEKLRSLTSFLAEEKKRMQONIQELEQEEESAR	298	Qy 253 --TASDQESGDFEL----NRLKEENEKLRSLTSFLAEEKKRMQONIQELEQEEESAR 298
Db	905	RUTAKKQEL--EPICHIDLEARVEEECRQHQI--AEKKRMQONIQELEQEEESAR	959	Db 905 RUTAKKQEL--EPICHIDLEARVEEECRQHQI--AEKKRMQONIQELEQEEESAR 959
Qy	299	QEL-----VERIHSLEVARAVAEROPQWEEKETLLOQPKSMAQOLYREK----	346	Qy 299 QEL-----VERIHSLEVARAVAEROPQWEEKETLLOQPKSMAQOLYREK---- 346
Db	960	OKIQLQEVYTTAEKKLKEEDIVLDQNLKL--AKEKKLLEMRMSFTNLTEEEEKS KS	1017	Db 960 OKIQLQEVYTTAEKKLKEEDIVLDQNLKL--AKEKKLLEMRMSFTNLTEEEEKS KS 1017
Qy	347	-----VNA!QAOQYCELQERDQAYSAQDAOR--EISOSLVERDS--LRQRFVLTQVC	397	Qy 347 -----VNA!QAOQYCELQERDQAYSAQDAOR--EISOSLVERDS--LRQRFVLTQVC 397
Db	1018	LAKLKNKHNKHEMTDLEER----LRREEKQOELETRKLKGDDSDHQLQELQRIA	1072	Db 1018 LAKLKNKHNKHEMTDLEER----LRREEKQOELETRKLKGDDSDHQLQELQRIA 1072
Qy	398	EURTQLRQLQAEPGVLKQEARTEPCPREKQLVRAHICPRDDSDSLSVSTESOLS	457	Qy 398 EURTQLRQLQAEPGVLKQEARTEPCPREKQLVRAHICPRDDSDSLSVSTESOLS 457
Db	1073	EUKIQLSKKEEEQQLAQL--ARYVEEAQKMNALKK!REL-----ESQI-- 1113		Db 1073 EUKIQLSKKEEEQQLAQL--ARYVEEAQKMNALKK!REL-----ESQI-- 1113
Qy	458	DLSATSSRELVDSSRFSSSPAPPSSQSLXKRAVEDFGEPPWSSSCLEPEGDPGALGAK	517	Qy 458 DLSATSSRELVDSSRFSSSPAPPSSQSLXKRAVEDFGEPPWSSSCLEPEGDPGALGAK 517
Db	1114	---TELQDLESERASRNKAERQK-----RDLGELEAKTELEDTLDSTAQAQELR	1162	Db 1114 ---TELQDLESERASRNKAERQK-----RDLGELEAKTELEDTLDSTAQAQELR 1162
Qy	518	AGDPH---LDYELLDPLDL--PQLESSLQPVSPGRDLDVSEGVLMRPPA-----	562	Qy 518 AGDPH---LDYELLDPLDL--PQLESSLQPVSPGRDLDVSEGVLMRPPA----- 562
Db	1163	SKREQEYTVLKKTLEDEAKTHAQIEMRKQISQATELAEQLEOTKRVKANLEKAQAL	1222	Db 1163 SKREQEYTVLKKTLEDEAKTHAQIEMRKQISQATELAEQLEOTKRVKANLEKAQAL
Qy	563	--RRILSQVTMLAFOG-----DALLEQISVGGNLTGIFIRVTPGSAODMA	608	Qy 563 --RRILSQVTMLAFOG-----DALLEQISVGGNLTGIFIRVTPGSAODMA 608
Db	1223	ESRAELNEVKVLLQGKDAEHKRKVKDAQLQELQV-----KFTGE----	1265	Db 1223 ESRAELNEVKVLLQGKDAEHKRKVKDAQLQELQV-----KFTGE---- 1265
Qy	609	LRPGTQTYMDVEASEPLFKAVLEDTTLLEEAVGLRRLWDGFCCLSVKYNTD-GYKRLQ	666	Qy 609 LRPGTQTYMDVEASEPLFKAVLEDTTLLEEAVGLRRLWDGFCCLSVKYNTD-GYKRLQ 666
Db	1266	-----RVTKLSTAERNLKVLOVE--LDNVYGLLNGQSDS--KSRIKAKDFSALESQQL	1311	Db 1266 -----RVTKLSTAERNLKVLOVE--LDNVYGLLNGQSDS--KSRIKAKDFSALESQQL 1311
Qy	667	DLEAKV-----ATSGDSFYIRVNLAMEGRAKGELQ--VHCNEVLHVTD	707	Qy 667 DLEAKV-----ATSGDSFYIRVNLAMEGRAKGELQ--VHCNEVLHVTD 707
Db	1312	DQELLQOETRKLKFSTKLQTEDEKNAKLEQEEEEBAERNLEQISVLLQQAQAVEARK	1371	Db 1312 DQELLQOETRKLKFSTKLQTEDEKNAKLEQEEEEBAERNLEQISVLLQQAQAVEARK 1371

RESULTS

A;Sequence number: A33977; MUID:90046668
A;Accession number: A33977
A;Molecule type: mRNA
A;Residues: 1-1959
A;Cross-references: GB:M26510; PID:9212382; PID:AAA48974..1; PID:g212383;
R;Katsuragi, Y.; Yamagisawa, M.; Inoue, A.; Masaki, T.
A;Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in cells
A;Reference number: S06116; MUID:90032648
A;Accession number: S06116
A>Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1959
A;Cross-references: GB:X17589
A;Note: this translation is not annotated in GenBank entry GMHCFCMHA, release 114
R;Hodge, T.P.; Cross, R.; Kendrick-Jones, J.
J. Cell Biol. 118, 1085-1095, 1992
A;Title: Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
A;Reference number: A33977
A;Accession number: A33977
A;Molecule type: mRNA
A;Residues: 1-1959
A;Cross-references: GB:M26510; PID:9212382; PID:AAA48974..1; PID:g212383;
R;Katsuragi, Y.; Yamagisawa, M.; Inoue, A.; Masaki, T.
A;Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in cells
A;Reference number: S06116
A;Accession number: S06116
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 716-1008 <KAT>
A;Cross-references: GB:X17589
A;Note: this translation is not annotated in GenBank entry GMHCFCMHA, release 114
R;Hodge, T.P.; Cross, R.; Kendrick-Jones, J.
J. Cell Biol. 118, 1085-1095, 1992
A;Title: Role of the COOH-terminal nonhelical tailpiece in the assembly of a vertebrate
A;Accession number: A43422; MUID:92381096
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1900-1959 <HOD>
A;Experimental source: brush border
A;Note: sequence extracted from NCBI backbone (NCBIB:111947)
C;Supfamily: myosin heavy chain; myosin motor domain homology

Qy	708	TMFQOCGCVH - AHVNSYSTMKD - - - - -	-TAHGTI - PNSRAQQQLALIQDMT -	753
Db	1372	KMDGLGCLEIAEEKKLQKDLESLTQRYEEXIAAYDKLETKTRQLQQEDDIAYVLDLHD	1431	
Qy	754	QCQTRVKPSSGGPQKLVRSMKAKASPLRSFQDPSRMEGSSTCFWAESCLTL	813	
Db	1432	QRQFVSNLEKK - - QKFDDOLAAEKNI SAKYA EERDAAEAREKE	-- - - - -	TKAISL 1481
Qy	814	VPTVLWPHRPARPVLLVPR A - - - - -	-VGKILSEKLCLLQGF	851
Db	1482	- - - - - ARALLEATEQKAELERYNKQFRTEMEDIMSSKDDVGKVHE	-- - - - -	-LEKA 1526
Qy	852	KKCCLAESLSQ - - - EYEBAWSORGDLIOQEGETSGGRCWYTRIAV	-- - - - -	ESL 894
Db	1527	KRALDQYBEMKTOLEELE - - - - -	-DEQATEPAKLREUNQMAQAFDRDLGIRDQN	1580
Qy	895	MEKNTHAL - - - LDVQD L - - - SVCTLHR - - - MDIFPIVIVHSVNEK -	- - - - -	930
Db	1581	EERKQLITQVREMEVLEDELERKORSIAVAAREKKLELDKLDESHIDTANKNRDEAIKHV	1640	
Qy	931	- - - - - MAKKKKGKLQRGTSQEPOLLEARQEGDL	960	
Db	1641	RKLOAQMDYMRLEUDTRTSRSEETLAQOKENEFKU	1675	

search completed: June 13, 2002, 09:24:56
Job time: 154 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: June 13, 2002 09:19:57 ; Search time 39.24 Seconds

(without alignments)

2841.949 Million cell updates/sec

Perfect score: 5149

Sequence: 1 MGELCRRDSALTALDEETLW VRQIAADEQKKVVTEOSPR 1004

scoring table: BLOSUM62

Gapext 0.5

Searched: 74754 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802:*

1: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*

21: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*

22: /SIDSL/gcadata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	5149	100.0	1004	AEE07164
2	5034	97.8	1139	AEE07165
3	1239.5	24.1	1147	Human caspase recr
4	887.5	17.2	1032	Human caspase recr
5	566.5	11.0	746	Human caspase recr
6	512.5	10.0	536	Rat caspase recr
7	493	9.6	536	Human caspase recr
8	396.5	7.7	366	Human protein sequ
9	309	6.0	350	Human polypeptide
10	282	5.5	1445	Drosophila melanog
11	262	5.1	2101	Sequence of the in

DT

XX

DE

XX

AEE07164

ID

AEE07164

AC

AEE07164;

AC

AEE07165

ID

AEE07165

AC

AEE07165;

AC

AEE07166

ID

AEE07166

AC

AEE07166;

AC

AEE07167

ID

AEE07167

AC

AEE07167;

AC

AEE07168

ID

AEE07168

AC

AEE07168;

AC

AEE07169

ID

AEE07169

AC

AEE07169;

AC

AEE07170

ID

AEE07170

AC

AEE07170;

AC

AEE07171

ID

AEE07171

AC

AEE07171;

AC

AC

AC

AC

AC

AC

AC

FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 91..93
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 114..117
 FT Modified-site /note= "N-glycosylation site"
 FT 117..122
 FT Modified-site /note= "N-myristoylation site"
 FT 121..123
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT Domain 126..420
 FT Modified-site /label= Coiled_Coil_domain
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 130..135
 FT Modified-site /note= "N-myristoylation site"
 FT 134..137
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 161..166
 FT Modified-site /note= "N-myristoylation site"
 FT 165..168
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 220..227
 FT Modified-site /note= "Tyrosine kinase phosphorylation site"
 FT 221..224
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 239..325
 FT Domain /label= k_Box_domain
 FT Modified-site 240..243
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 250..252
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 253..256
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 259..262
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 260..263
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 290..293
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 297..300
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 307..309
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 307..310
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 355..365
 FT Modified-site /note= "Tyrosine kinase phosphorylation site"
 FT 366..368
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 366..369
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 378..381
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 384..386
 FT Region /note= "Protein kinase C phosphorylation site"
 FT 385..406
 FT Modified-site /note= "Leucine zipper pattern"
 FT 449..452
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 463..466
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 463..465
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 470..472
 FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 501..504
 FT Modified-site /note= "Casein kinase II phosphorylation site"
 FT 511..516
 FT Domain /label= SH3_domain
 FT Modified-site /note= "N-myristoylation site"
 FT 568..660
 FT Modified-site /label= PDZ_domain
 FT 587..592
 FT Modified-site /note= "N-myristoylation site"
 FT 589..592
 FT Modified-site /note= "N-glycosylation site"

FT Modified-site 602..605 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 634..637 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 653..655 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 674..677 /note= "Casein kinase II phosphorylation site"
 FT Domain 676..745 /label= SH3_domain
 FT Modified-site 714..719 /note= "N-myristoylation site"
 FT Modified-site 725..727 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 725..728 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 733..738 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 737..740 /note= "N-glycosylation site"
 FT Modified-site 759..761 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 760..763 /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
 FT Peptide 785..793 /note= "Peroxisomal targeting signal"
 FT Modified-site 796..799 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 800..805 /note= "N-myristoylation site"
 FT Domain 826..1004 /label= Guanylate_kinase_domain
 FT Modified-site 842..844 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 860..863 /note= "Casein kinase II phosphorylation site"
 FT Domain 868..370 /note= "Protein kinase C phosphorylation site"
 FT Region 870..872 /note= "RGD cell attachment sequence"
 FT Modified-site 893..896 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 926..929 /note= "Casein kinase II phosphorylation site"
 FT Peptide 941..949 /note= "Peroxisomal targeting signal"
 FT Modified-site 944..947 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 976..979 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 980..985 /note= "N-myristoylation site"
 FT Modified-site 1002..1004 /note= "Protein kinase C phosphorylation site"
 XX WO200159065-A2.

PN XX
 XX PD 16-AUG-2001.
 XX PF 22-JAN-2001; 2001WO-US2007.
 XX PR 09-FEB-2000; 2000US-0181159.
 XX PA (MILL-) MILLENNIUM PHARM INC.
 XX PI Bertin J;
 XX DR WPI; 2001-497073/54.
 XX N-PDB; AAD13447.
 XX PT An isolated caspase recruitment domain polypeptide useful for regulating growth and cell death and useful for the treatment of cancer

PT	vnsytmkdtaaigtipysraaggialigdmtqgctvtrkssggqkivtrvsmdkak	780
PS	QY 721 ASPLRISFDQRQLDPSMEGSSTCFWAECLTVPLVPPARPPVLLPRAVKI 840	
XX	Db 781 asplrlsfdqrqlldpsmegsstcfwaecltvplvpparppvllprrgkki 840	
CC	QY 841 LSEKLCLQGFKRCLAYLSQBERYEANSQRGDLIQEGEVSGRCWYRHAESLMRNT 900	
CC	Db 841 lsekllcqfkkclaylsqeeeyearsgrdiqeegsrgcrwvtrhaveslmrnt 900	
CC	QY 901 ALIDVOOLDSVCTLHRMDIFPTIVTHSYNEKMAKKLKGLQRIGTSECOLLEARQEGLD 960	
CC	Db 901 alidvqldsvctlhrmdifptivthsynekmaakkllqrgtseegleareqegl 960	
CC	QY 961 DRAPCLYSSLADPDGWSLDGLLSCVRRAIADEOKKVWTEQSPR 1004	
CC	Db 961 drapcllysslapdgdwsldgliscvrriaadeqjkvvwtqsp 1004	
SQ	vnsytmkdtaaigtipysraaggialigdmtqgctvtrkssggqkivtrvsmdkak	
XX	Sequence 1004 AA;	
Query Match	100.0%; Score 5149; DB 22; Length 1004;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MGECLCRDSALTALEDFTLWEMMESHRRHTRVRCIPSRLTPTYLROAKVLCOLDEEEVLYHS 60	
Db	1 mgeclcrdsalataledftlwemmeshrrhtrvrcipsrltptylroakvlcoldeeevlyhs 60	
QY	61 PRLTNSAMRAGHLLDLKLTKRGNGKAIAFLSLEKFKHNPDPDVTLVTGLQPDVFNSNTGGLME 120	
Db	61 prltksamraghlldlkltrkgngkaiaflslekhnpdpdyvtlvtglqpdvfnsntgglme 120	
QY	121 TSKRLTECLAGAIGLSQEBELNQERKGKEFVLRRCQQLQEHHLGAETRAEGLIQLOLEADHSMR 180	
Db	121 tskrlteclagaisglseelngkqkevllrrcqqlqehhlgaetraeqlghleadhsrm 180	
QY	181 KREVSAAHFHEVLRKDEMMLSLHYSNALQKEELAASRCSRCSQEEYLKLQBLQRANMVS 240	
Db	181 krevsaahfhevirlkdemmlslhyssnalqkeelaasrccsqeeyylkqeqlqransmvs 240	
QY	241 SCELLEQFSLRTASDOESGDEELNLKEENKEKSLTSLAEKDITLEQSDEARGSRQE 300	
Db	241 scelleqfslrtsaqesgdeelnlkeeneksltslaekdileqsldeargsrqe 300	
QY	301 LVERIHSLREAVAAEQRQEWEEKEQTLLQFQSKIMACQLYREKVNLAQVCLEQKE 360	
Db	301 lverinsrlreavaaerqrqewyweekeqtlfqfqskimacqlyrekvnlaqvceqke 360	
QY	361 RDQAYSARDSAOREISOSLVYKDSLRRQVFELTDQVCELRQLQAEPPGVLKOEART 420	
Db	361 rdqaysardsaoreisoslsvyksdlrrqvfeltdqvcrlqrlqaeppgvlkoeart 420	
QY	421 REPCPREKQRLVRMHATCPRDSDCSLVSSTESQQLSLSTATSSRELVDSPRSSPAPPS 480	
Db	421 repcprekqrvmhaticprdsdcslvsstesqqlslstatssrelvdsprrsspapps 480	
QY	481 QOSLYKRVAAEDFGEEPWMSFSCLEIPEGDPEALPGAKAGDHDYELLDTPDLPQLESSL 540	
Db	481 qoslykrvaaedfggepwmsfscleipegdpealpgakagdhdyelldtpdlpqlessl 540	
QY	541 QPVSPGRUDVSESGVLMRRARRILSQVMTLAFOQDALLQISVIGGNLFGIFTHRVTP 600	
Db	541 qpvspgrudvsesgvlmrrarrilsqvmtlafoqdallqisviggnlfgifhrvtp 600	
QY	601 GSAADOMALRPGTQYMDYASESEPLFKAVLEDTLEAVGLLRRVDFGCCSVKVNTDG 660	
Db	601 gsaadomalrpqtqymdyaseseplfkavledtelleavllrrvdfgccsvkvtndg 660	
QY	661 YKRLQDLEAKVATSGDSFYRVNLAMEGRKGELQHVNCENLVHTDTMFQGCGWHAH 720	
Db	661 ykrllqdleakvatsgdsfyrvnlamegrkgelqhvncenlvhtdtmfqgcfgewhah 720	
QY	721 VNSYTMKDTAAHGTIPNYSRQQQLIALIQDMTQQCTVTRKPSSGGPQKLVTRIVSMDKAK 780	
PS	Disclosure: Fig 2A-2C; 109pp; English.	
XX	The present sequence is predicted human caspase recruitment domain 14 (CARD-14). The CARD-14 gene is located on chromosome 17. The CARD-14 is used for the detection of modulators that modulates the ability of CARD-14 to bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation of nuclear factor-kappa B (NF- κ B). The CARD-14 is useful for regulating growth and cell death and useful for the treatment of cancer. It is also useful for the treatment of autoimmune disorders (e.g., systemic lupus erythematosus), neurological disorders (e.g., Alzheimer's and Parkinson's disease, inflammatory disorders, haematological disorders (e.g., anaemia, myelodysplastic syndromes), myocardial infarctions, strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis), cell signalling disorders: cytosolic immunosuppressive; nootropics; neuroprotective; antiviral; antibacterial.	
CC	Human predicted caspase recruitment domain 14 (CARD-14).	
XX	ID AAE07165 standard; Protein: 1139 AA.	
AC	AAE07165;	
XX	AAE07165;	
DT	06-NOV-2001 (first entry)	
XX		
DE		
XX		
KW	Human; caspase recruitment domain-14; CARD-14; chromosomal disorder; cell growth; cell death; cancer; therapy; neurological disorder; systemic lupus erythematosus; neurologica	
KW	autoimmune disorder; systemic lupus erythematosus; neurologica	
KW	Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia; hematological disorder; myelodysplastic syndrome; myocardial infarction; stroke; immune disorder; Crohn's disease; allergic rhinitis; infection; cell signalling disorder; cytosolic immunosuppressive; nootropics; neuroprotective; antiviral; antibacterial.	
XX		
OS	Homo sapiens.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 700	
FT	/note= "Encoded by TGG"	
XX		
PN	WO200159065-A2.	
XX		
PD	16-AUG-2001.	
XX		
PF	22-JAN-2001; 2001WO-US02087.	
XX		
PR	09-FEB-2000; 2000US-0181159.	
XX		
DR	(MILL-) MILLENNIUM PHARM INC.	
PA	DR ; 2001-491073/54.	
XX		
PI	PA ; AAD13448.	
XX		
Bertin J;	An isolated caspase recruitment domain polypeptide useful for regulating growth and cell death and useful for the treatment of cancer	
PT	-	
PT	-	
PT	-	
XX		
PS	Disclosure: Fig 2A-2C; 109pp; English.	
XX		
CC	The present sequence is predicted human caspase recruitment domain 14 (CARD-14). The CARD-14 gene is located on chromosome 17. The CARD-14 is used for the detection of modulators that modulates the ability of CARD-14 to bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation of nuclear factor-kappa B (NF- κ B). The CARD-14 is useful for regulating growth and cell death and useful for the treatment of cancer. It is also useful for the treatment of autoimmune disorders (e.g.,	

CC systemic lupus erythematosus), neurological disorders e.g., Alzheimer's disease, inflammatory disorders, haematological disorders (e.g., anaemia, myelodysplastic syndromes), myocardial infarctions, CC strokes, immune disorders (e.g., Crohn's disease, allergic rhinitis), cell signalling disorders and certain viral and bacterial infections.

XX Sequence 1139 AA;

Query Match 97.8% Score 5034; DB 22; Length 1139;
Best Local Similarity 95.1%; Pred. No.: 0; Mismatches 0; Indels 50; Gaps 3;
Matches 995; Conservative 0;

Qy 1 MGEICRDRSALTADEETLWEMAMESHHRHTRVRCICPSRLRQAKVLQDDEEVLHS 60
Db 1 mgelcrdrsaltaaldeetlwemameshhrivrcicpsrlrqrqkvlqlqdeevelhs 60
Qy 61 PRLTNSAMRAGHLLDKTRGKNGATAFLESISKFHNPDVVTLYTGTQDVSFS --- 116
Db 61 prltnsamraghlldktrgkngataflesikfhnpdvvtlytqdpdvsfs 120
Qy 117 -----GLMTESKLITECLAGAIGSQEEENIQEKGKEYVLLRRCQLOE 158
Db 121 dfdqlagtsrnrlrlvttgqimetskiteclagaigsqeelinrqekgvllrrccqlo 180
Qy 159 HGLAETRAEGLAQLEADHSRMKREYSAHFEVILRKIDEMLSLHYSNALQEKELAASR 218
Db 181 hglraetraeglhqleahdsrnkrevahfvrlkdemlslhyasnlaqekelaasr 240
Qy 219 CRLSQEELYLKLQELQLQRANMVSSCELEQEQSLRTASDQESSDEELNLKENEKRSLT 278
Db 241 crsiqueelylkqelqgranmvssceleqeqslrtasdqessdeelnlkenekeirs 300
Qy 279 FSLAEKDITLEQSDLEARSQELIVERTHSLRFAVAATRQE ----- 320
Db 301 fslaekdileqsdearsqeliverhsleravaerqrqeqarpsseellftvhvhs 360
Qy 321 QYWEEKEQTLLQFOKSQWACQOLREKYKNALQAOYCQCELOKERDOAYSARDSAQCETISOLV 380
Db 361 qyweeqteqtllqfoksqmacqlyrekvnalqaoycqceolkerdoaysardsaqcetisolv 420
Qy 381 EKDSLRQQYFEITDQVCELRTOQLRQLQAEPGPVVKOEARTRCPREPKQRLVRMHAICPR 440
Db 421 ekdsirrgyfeitdqvceirtqrqlqaepgpvylqkqartrepqrlvrmlhaicpr 480
Qy 441 DDSCLVSYSTESLSDISATSRELYDSFRSSSPAPPSQQLYKRAEDGEEPPSFS 500
Db 481 ddsclslystsseqqlsdlsatsrsrelydsfrssspapssqqlslykrvaedggeppfs 540
Qy 501 SCLETEPGDGPALFGAKAGDPHDYELLDTADLPQLESSLQPVSPGRLDVSB ----- 552
Db 541 scletpgdpalfgakagdpdyelldtadlpqlesslqpssqqlslykrvaedggeppfs 600
Qy 553 -SGYLMRRRPARRLSQLTYMLAFOGDALLEQIVGGNLTGIFTIHRVTPGSAADONMALR 610
Db 601 acsgymrrrparrlsqltymlaqgdalleqivggnltgiffihrvtpgsadqmair 660
Qy 611 PGTOQWWDUEASEBPLFKAVLDTLEEVANGLIRRVDQFCCLSYKVNTDGKYLQLEA 670
Db 661 pgtqivmveaseplfkavledtittleavqglrrvdgfcclsvkvntdgkylqlea 720
Qy 671 KVATSGDSFYTRVNIALEGRAKAGEQVHNELVHTDIFMFGGCCWHAHRVNSYTMKDTA 730
Db 721 kvatsgdsfytrvniamegrakeqvhnehvtdifmgccwahrvnsytmdta 780
Qy 731 AHGTIPNYSRAQQQLALIOMTOQCTYTRKPSSGGPQKLVTSMDKAKASPLRLSFDR 790
Db 781 ahgtipnysraqqqlalqmdtgctvkrkpssggpqlvrvsmdkakasplrlsfdr 840
Qy 791 QOLDPSRMEGSSTCWEAFSOLUINPYTLMWPHRPARPREPVLYPRAVKILSEKULQG 850
Db 841 qqlqpsrmegsstcweafsoluinpytlwmphrparprpvlyptraekilcqg 900

CC	Qy	851 FKKCLAEYLSQEYFAWSORGDIQEOGEVSGGRCMVTRHVESLMEKNTHALLDVQLDSV 910
CC	Db	901 fkkclaeylsqeeyfaawsgqdiigeveggcvcvtravesmeknthalldqidsv 960
CC	Qy	911 CTLHMRDIFPIVIVHVSYNEKMAKKLKKLQLRGTSEEQLEAARQEEGIDRACPLOSS 970
CC	Db	961 ctlhmrdfpiivivhvsynekmaikkqlqlgtseeqleaarqeegidrapclys 1020
XX	Qy	971 APDGWSDLGILSCVRQIADEQKKV 996
XX	Db	1021 apdgwsdlgilsrvrqiaadeqkkv 1046
SQ		
	RESULT 3	
	AAU01207	Human caspase recruitment domain, CARD-11 polypeptide.
	ID AAU01207	standard; Protein: 1147 AA.
	XX	
	AC AAU01207;	
	XX	
	DT 12-SEP-2001	(first entry)
	XX	
	DE Human caspase recruitment domain, CARD-11 polypeptide.	
	XX	
	FH Key	Location/Qualifiers
	FT Domain	6..112 "CARD domain"
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 7..9
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 7..10
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 100..102
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 100..103
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 105..107
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 106..109
	FT Modified-site	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site" 130..143
	FT Domain	/note= "Coiled coil domain"
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 162..165
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 168..171
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 175..178
	FT Modified-site	/note= "Tyrosine kinase phosphorylation site" 182..185
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 189..195
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 282..285
	FT Modified-site	/note= "N-glycosylation site" 286..289
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 243..245
	FT Modified-site	/note= "Protein kinase C phosphorylation site" 378..381
	FT Modified-site	/note= "Casein kinase II phosphorylation site" 429..432
	FT Modified-site	/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site" 459..461

FT Modified-site /note= "Protein kinase C phosphorylation site"
 FT 471..474 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 472..475 /note= "N-glycosylation site"
 FT Modified-site 476..479 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 508..510 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 510..513 /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
 FT Modified-site 558..560 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 578..581 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 584..587 /note= "N-glycosylation site"
 FT Modified-site 587..592 /note= "N-myristoylation site"
 FT Modified-site 634..637 /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
 FT Domain 635..748 /note= "PDZ domain"
 FT Domain 635..1147 /note= "MAGUK domain"
 FT Modified-site 638..641 /note= "Glycosaminoglycan attachment site"
 FT Modified-site 678..683 /note= "N-myristoylation site"
 FT Modified-site 687..689 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 692..695 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 698..703 /note= "N-myristoylation site"
 FT Modified-site 710..715 /note= "N-myristoylation site"
 FT Modified-site 725..728 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 761..766 /note= "N-myristoylation site"
 FT Modified-site 764..767 /note= "Casein kinase II phosphorylation site"
 FT Domain 766..834 /note= "SH3 domain"
 FT Modified-site 776..779 /note= "N-glycosylation site"
 FT Modified-site 779..782 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 787..789 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 816..819 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 823..828 /note= "N-myristoylation site"
 FT Modified-site 847..850 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 853..858 /note= "N-myristoylation site"
 FT Modified-site 857..859 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 872..875 /note= "Casein kinase II phosphorylation site"
 FT Domain 882..1147 /note= "Guanylate kinase (GUK) domain"
 FT Modified-site 897..900 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 917..922 /note= "N-myristoylation site"
 FT Modified-site 926..929 /note= "Casein kinase II phosphorylation site"

FT Modified-site 935..937 /note= "Protein kinase C phosphorylation site"
 FT Modified-site 1003..1006 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 1010..1018 /note= "Tyrosine kinase phosphorylation site"
 FT Modified-site 1050..1055 /note= "N-myristoylation site"
 FT Modified-site 1088..1091 /note= "Casein kinase II phosphorylation site"
 FT Modified-site 1120..1123 /note= "Casein kinase II phosphorylation site"
 PN WO200140468-A2.

XX PD 07-JUN-2001.
 XX PF 01-DEC-2000; 2000WO-US32716.
 XX PR 03-DEC-1999; 99US-0168780.
 XX PR 18-FEB-2000; 2000US-0507531.
 XX PR 25-FEB-2000; 2000US-0513904.
 XX PR 10-OCT-2000; 2000US-0685791.
 XX PA (MILL-) MILLENNIUM PHARM INC.
 XX PI Berntin J;
 XX DR WPI; 2001-367809/38.
 XX DR N-PSDB; AAS05389.
 XX PT Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10, CARD-11, useful as targets for therapy, as immunogens, and in screening and detection assays -
 XX PS Claim 9; Fig 14A-14C; 145pp; English.
 XX CC The present sequence represents novel human caspase recruitment domain, CARD-11. The polynucleotide encoding this sequence was isolated from a human T-cell cDNA library. Also described are novel human sequences for CARD-9 and CARD-10 (AAU01205, AAU01206) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and CARD-11 interact with Bcl-10 which is thought to activate nuclear factor (NF)-kappa B and apoptosis. The sequences of the invention can be used for treating a disorder associated with abnormal levels of apoptosis by modulating the expression or activity of CARD-9, CARD-10, or CARD-11. They can be used for the treatment of hyperproliferative disorders (e.g. systemic lupus erythematosus), inflammatory disorders (e.g. cancer), autoimmune disorders (e.g. Alzheimer's disease), neurological disorders (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD polypeptide, polynucleotide and an antibody which selectively binds to CARD can be used in screening and detection assays (e.g. chromosomal mapping, tissue typing), predictive medicine (prognostic assays, monitoring clinical trials, and therapy (treatment and prophylaxis). The CARD polypeptide may be used to screen for drugs that bind to and/or modulate it. CARD sequences are potential targets for regulating inflammation, cancer, NF-kappa B signalling, stress-related response and encoding CARD can be used to create transgenic animals.
 XX SQ Sequence 1147 AA;
 XX Q Query Match 24..1%; Score 1239..5; DB 22; Length 1147;
 XX Best Local Similarity 30..5%; Pred. No. 7..1-e-93;
 XX Matches 354; Conservative 204; Mismatches 402; Indels 201; Gaps 33;
 QY 15 DEETLWEMMESEHRHRRVRCIGPSRLTPYLROAKVLQDLDEREVLYHSRSLTNSAMRAGHLL 74
 DB 11 :|||::|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 QY 75 DILKTRKGNGATAFLESSLKFRNPDPVYLTGQPVDFSNFSGLMETSKLDECLAGAIGS 134

Db	71	dilhtkqrgyyvffleslefypelykvtgkeptrrfstivveegheglthfilmnevik	130		RESULT 4	AAU01206
Qy	135	LOEELNQERGKQEVLDRRCQQIQ-EHGLAETRABSLHQLEADHSRMKREVAHPEVLR	193		ID AAU01206 standard; Protein; 1032 AA.	
Db	131	lqqmvalqdklqrceilarllrqdtkqmtlrv-eiltfqerykmkeerdsyndelvk	189		XX	AAU01206;
Qy	194	LDEMULSLSYNSALOEKELAASCRSRLQEEFLQMLQRANMVSSECELEQFQLRT	253		AC	XX
Db	190	vddnlyiamryaqiseekmavmvsrdqleidqkhrlnk-meeckle-rnqslikl	246		DT	12-SEP-2001 (first entry)
Qy	254	ASDQES-GDEBLNLKEENEKLRSLSFSL-----AEKDILEQSLDEAR	295		XX	Human caspase recruitment domain, CARD-10 polypeptide.
Db	247	kndieropkkekqvlereenmlktknqelqsiqagkrslipdskaillehrkeal	306		DE	XX
Qy	296	GSRQELVERIISLRLREVAABAERQREOWEEKETQTLQFOKSMAOLYREKVNALQAQVC	355		KW	Human; caspase recruitment domain; CARD-10; Bcl-10; NF-kappaB; apoptosis; hyperproliferative disorder; autoimmune; neurological; inflammatory disorder; viral infection; stress-related response.
Db	307	edqrqlvnrlyqearqeakrkyleekdeklkcst1qkdemymkrmntmlqle	366		KW	XX
Qy	356	BLOCKERDQAYSARDSAQREISQELVERKDSLRRQVFELTDO-----VCELR	401		OS	Homo sapiens.
Db	367	evererdqafansirdeaqtqyqscleakdkykgireeakndemriemvrreacivnles	426		XX	
Qy	402	QRLQQAE-----PGVILQ-----EARTRCPGPREAK-----	428		Key	Location/Qualifiers
Db	427	kirrlksdsmnlqslprnlptvtsqdffgasprtngeadssseespedshyf1py	486		FT	15..20 /note= "N-myristoylation site"
Qy	429	--QRLVRMHAL-CPRDDSDCPLSVSSTESQLL-----SDLSATSS-SRELVD	476		FT	FT /note= "Casein kinase II phosphorylation site"
Db	487	hpqqrmnlkjqlqrakspslkrttsdfqakgheseegtdaspssccsgs1ptntsftkmqp	546		FT	FT /note= "CARD domain"
Qy	477	--APPSQSLVYKRAADDGEFPWSFSCSLETPGDPGALPAGKA-GDPHL	523		FT	FT /note= "Protein kinase C phosphorylation site"
Db	547	prsrssimtsiteappngndsvirykedaphr----stevedndsgqfdalldddshe	600		FT	FT /note= "Amidation site"
Qy	524	DY-----ELIDTADLPQLESSLQPV-----PGRLDYSESGYLMRRPARR	564		FT	FT /note= "Casein kinase II phosphorylation site"
Db	601	rystfgpsihsssshhsgsegldaydlevqn1mfrrkslerfrpsrtsvhvrqpgps--	658		FT	FT /note= "N-myristoylation site"
Qy	565	IISQVPMIAFQCDALLEQFVQIGVGGNLTGIFTHRVTPGSAADOMALRPGTQIVMVYEASE	624		FT	FT /note= "Coiled coil domain"
Db	659	---vqhttingdsllsqitlggnarngrsfhvsvpgslaekagareghqlllegcigr	714		FT	FT /note= "Tyrosine kinase phosphorylation site"
Qy	625	PFLKAVLEDITLLEAVGLLRRVDFGCCSVKVNNTDGKYLQDLEAKVATSGDSFYIRVN	684		Region	230..251
Db	715	ergsvpldtctkeawtqrsqspvlykvnehyrklykdmeglltsqsfirlr774	774		FT	FT /note= "Leucine zipper homology region"
Qy	685	LAMEGRAKG-EIQUVHCNEDEVLIWTTDMFQGCCWHAIRVNSYTMKDTAAHGTPINYSRAQQ	743		FT	FT /note= "Casein kinase II phosphorylation site"
Db	775	Inrissqldactaslikcdvvnvrdtnyqdrhewpcarvpitfdhd1dm-gtppsysraqq	833		FT	FT /note= "Casein kinase II phosphorylation site"
Qy	744	QJJALIQTDMTOQ-----CTVTRKP-----SSGGPO-----N-----KLV	771		FT	FT /note= "Tropomyosin domain"
Db	834	llivkqgrlmhgsreedgthtlralntlqpeasaltsdprsprlsasflfgql1	893		FT	FT /note= "Protein kinase C phosphorylation site"
Qy	772	RIVSMKDKAK-----ASPL-----RLSFD-----RGQLDPSRMEGSSTCF	805		FT	FT /note= "Casein kinase II phosphorylation site"
Db	894	q1vsrsenkrmnsnervriispgs1arsidatktltkqelpelelgkn---	950		Domain	331..398
Qy	863	EYEAWSGORDIT-QEGEVSGGRCWYTRHAYESLMERKNTHALDVOLDSTVLHMRDIFP	920		FT	FT /note= "Casein kinase II phosphorylation site"
Db	1006	eifirqrktetiysreknphnafec-lapanjeavaaknhcileagtgtrdlkiqniyp	1064		FT	FT /note= "MAGUK domain"
Qy	921	IVTHSYNEKMKAKKKLGQLRGTSEEQLEMARQEGDLDRAPCLYSSLAPDGWSLDG	980		FT	FT /note= "N-glycosylation site"
Db	1065	lvfirvceknlkrfrkllprpet-eeffrvcrkelealpciyatvpmwgsvee	1123		FT	FT /note= "Casein kinase II phosphorylation site"
Qy	981	LSCVROIAIADOKKVWTEQ	1001		FT	FT /note= "N-myristoylation site"
Db	1124	lirrvkdkigeeqrktiwde	1144		FT	FT /note= "Casein kinase II phosphorylation site"

PT	Modified-site	512..514 /note= "Protein kinase C phosphorylation site"	PF 01-DEC-2000; 2000WO-US32716.
PT	Modified-site	549..552 /note= "Casein kinase II phosphorylation site"	XX PR 03-DEC-1999; 99US-0168780.
PT	Modified-site	558..560 /note= "Protein kinase C phosphorylation site"	PR 18-FEB-2000; 2000US-0507533.
PT	Modified-site	565..570 /note= "N-myristoylation site"	PR 25-FEB-2000; 2000US-0513904.
PT	Modified-site	570..573 /note= "Casein kinase II phosphorylation site"	PR 10-OCT-2000; 2000US-0685791.
PT	Modified-site	595..598 /note= "N-glycosylation site"	XX PA (MILL-) MILLENIUM PHARM INC.
PT	Modified-site	603..605 /note= "Protein kinase C phosphorylation site"	XX PA Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10, CARD-11, useful as targets for therapy, as immunogens, and in screening and detection assays -
PT	Modified-site	638..641 /note= "Glycosaminoglycan attachment site"	PT PT 2001-367809/38.
PT	Modified-site	641..644 /note= "Protein kinase C phosphorylation site"	DR DR N-PSDB; AAS05388.
PT	Modified-site	656..661 /note= "N-myristoylation site"	XX PS Claim 9; Fig 10A-10C; 145PP; English.
PT	Modified-site	681..684 /note= "SH3 domain"	CC The present sequence represents novel human caspase recruitment domain, CARD-10. The polynucleotide encoding this sequence was isolated from a human skin cDNA library. Also described are novel human sequences for CARD-9 and CARD-11. (AAU01205, AAU01207) and rat CARD-9 (AAU01204).
PT	Modified-site	690..693 /note= "casein kinase II phosphorylation site"	CC CARD-9, CARD-10 and CARD-11 interact with Bcl-10 which is thought to activate nuclear factor (NF)-kappaB and apoptosis. The sequences of the CC invention can be used for treating a disorder associated with abnormal CC levels of apoptosis by modulating the expression or activity of CARD-9, CC CARD-10, or CARD-11. They can be used for the treatment of CC hyperproliferative disorders (e.g. cancer), autoimmune disorders (e.g. CC systemic lupus erythematosus), neurological disorders (e.g. Crohn's disease), CC inflammatory disorders (e.g. Alzheimer's disease), and viral infection (e.g. HIV). The CARD Polypeptide, polynucleotide, CC and an antibody which selectively binds to CARD can be used in screening CC and detection assays (e.g. chromosomal mapping, tissue typing), CC predictive medicine (prognostic assays, monitoring clinical trials, and CC therapy (treatment and prophylaxis)). The CARD polypeptide may be used CC to screen for drugs that bind to and/or modulate it. CARD sequences are CC potential targets for regulating inflammation, cancer, NF-kappaB CC signalling, stress-related response and apoptosis in human disease. A CC host cell containing a polynucleotide encoding CARD can be used to CC create transgenic animals.
PT	Domain	704..772 /note= "SH3 domain"	XX Sequence 1032 AA;
PT	Modified-site	712..715 /note= "N-glycosylation site"	Query Match 17.2%; Score 887.5; DB 22; Length 1032;
PT	Modified-site	714..717 /note= "casein kinase II phosphorylation site"	Best Local Similarity 29.2%; Pred. No. 7.2e-64; Matches 317; Conservative 161; Mismatches 419; Indels 31;
PT	Modified-site	733..739 /note= "Tyrosine kinase phosphorylation site"	QY 15 DEETLWEMMESEHRHRITRCICPSRLTLYRQAKVLCOLDEEEVVLHSRDLTNSAMRAGHLL 74
PT	Modified-site	748..751 /note= "casein kinase II phosphorylation site"	Db 23 eedalwervegyhrllaralnpakltprylergervideqdeeeevlstryrfpcvrnrtgrml 82
PT	Modified-site	755..756 /note= "Protein kinase C phosphorylation site"	QY 75 DLJLKTRGKNGATAFLFLESILKFHNPDVYPLVTGLQPDVDFSNFSGLMERISKLTBCLAGIIGS 134
PT	Modified-site	754..757 /note= "Casein kinase II phosphorylation site"	Db 83 dlircrgkrgrgyafelelefyppehfcfltgapeaqcsmildeegeggglgfmltevrr 142
PT	Modified-site	761..766 /note= "N-myristoylation site"	QY 135 LQEELNQEKGQKEVLLRRCQOLOEHLGLAETRAEGLHQLEADHSRMKREVSAAHFEYVRL 194
PT	Modified-site	782..784 /note= "Protein kinase C phosphorylation site"	Db 143 lrearksgqlqrregqlqrgrvleeraglergrldrqgqarercqrlnredweagsleirrl 202
PT	Modified-site	809..814 /note= "N-myristoylation site"	QY 195 KDEMLSLSLHYNALQKEKELAASRCSRSLQEEYLKLQELQRANMVSSCELEIQQSRTA 254
PT	Domain	830..1032 /note= "N-myristoylation site"	Db 203 kdenymiamrlqgqseeksavirsrdqlqavdqlkvslr--leecal----rra 254
PT	Modified-site	830..832 /note= "Glyanilate kinase (GUK) domain"	QY 255 SDQESGDEELNRKKEENE-----KLRS----LTFSLAE-----
PT	Modified-site	868..870 /note= "Protein kinase C phosphorylation site"	Db 255 rgppgaaekekekekepdnvdleiraengqasirelqeqigqearspgapgs 314
PT	Modified-site	869..872 /note= "casein kinase II phosphorylation site"	QY 284 --KDILEQSLDEARGSRQELVERIHSLRERAVAAERREQWEEKBTLQFQKSXMAC 340
PT	Modified-site	882..885 /note= "Casein kinase II phosphorylation site"	Db 315 rillidlehdwreaqdsrqelcqkhlavqgelqwaellrdqylqemedlrkhtiqkdc 374
PT	Modified-site	893..898 /note= "N-myristoylation site"	XX WO20010468-A2.
PT	Modified-site	915..918 /note= "Anidation site"	XX PD 07-JUN-2001.
PT	Modified-site	947..949 /note= "Protein kinase C phosphorylation site"	XX
PT	Modified-site	981..986 /note= "N-myristoylation site"	XX
PT	Modified-site	1021..1026 /note= "N-myristoylation site"	XX
PT	Modified-site	1022..1024 /note= "Protein kinase C phosphorylation site"	XX
PT	Modified-site	1028..1031 /note= "Casein kinase II phosphorylation site"	XX

QY 341 QLYREKVNALQQVQVCELOKERDQAYASARDSAQREREISOLIVEKDLSRQQFELTDQVCELR 400
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST
 PA
 PI Ota T., Nishikawa T., Isogai T., Hayashi K., Ishii S., Kawai Y.;
 PI Wakamatsu A., Sugiyama T., Nagai K., Kojima S., Otsuki T., Koga H;
 XX
 DR WPT; 2001-524255/58.
 DR-NPSDB; AAK9-778.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 use in genetic manipulation -
 PT
 XX
 PS SEQ ID NO 3879; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesising full length cDNA
 CC molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesising the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 Sequence 746 AA;

Query Match 11.0%; Score 566.5; DB 22; Length 746;
 Best Local Similarity 28.5%; Preo. No. 1.e-37;
 Matches 223; Conservate 112; Mismatches 300; Indels 147; Gaps 26;

QY 285 DILEQSLDEARGSRQELYVERIHSRLRERAAAERQEQYWEKEQTLLQFOQSKMAMCOLYR 344
 PR 33 dilehdwreadqrdrqelcqklnavqge1lqwaesiirdy1qamedir1khrlqkdcddlyk 92

QY 345 EKVNALQVQVCELOKERDQAYASARDSAQREISOLIVEKDLSRQQFELTDQVCELRQLR 404
 PR 93 hrmavtlaqieeikerdqaqdrlqiqysisileddqy-kqvrgjeaerdeellltit 152

QY 405 QLOAEPPGVLKQBARTR_EPCPREKQVRIVRMHACIPRDSDPSLVSST----- 451
 PR 153 siegtkalllevqlqraqgtcl-----acasshs1cnslls1wslsefpsplgg 202

QY 452 -----ESQILSDLSATSSRELVDSSFR-----SSSPALPSQQLSYKRAEDFGEBP-WSFS 500
 PR 263 smadi----tgsv-----tktkwpgqlsssssssssssvwp1g 293

QY 561 PARRILSQTMLAFQGDALL-----EQTSIVGMNLTGIFIHRT-- 599
 PR 294 kpeglargcqdflnrsairvsgrsrspqggpepqdkopdgisfygdrwsavrrvrlsg 353

QY 600 PGSA----ADQMALRPGRQIVMDYEASBPLFKAVL-EDDTLEBEAVGLRRYDGFCCLL 653
 PR 354 pgarmepreqveag_eglacleaqag---rtllwngstips1-----mdskacs 404

QY 654 VKVNTDGYKRLQDLEAKVATSG-DSFYIVRNLMAGERAKGE-LQVHCVNEVLHVTDTMFQ 711
 PR 405 fh-----ealearwakpgaaefyviranitiperadhalcvkaeirivdsayk 454

QY 712 GGCOWHARVNSTMKDAAHGCTIPNYSRAQQQLIALIDWNOQCTVTRKPSS--GGQK 769
 PR 455 rrnewfcfrvdptlrd-lmgrtvpnytraqql---evqekcl---psxrhrsprs 504

RESULT 5
 ID AAM93822 standard; Protein: 746 AA.
 XX
 AC AAM93822;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3879.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.

Page 10

C be used to screen for drugs that bind to and/or modulate it. CARD
 C sequences are potential targets for regulating inflammation, cancer,
 C NF- κ appaB signalling, stress-related response and apoptosis in human
 C disease. A host cell containing a polynucleotide encoding CARD can be
 C used to create transgenic animals.

C	be used to screen for drugs that bind to and/or modulate it. CARD sequences are potential targets for regulating inflammation, cancer, NF- κ B signalling, stress-related response and apoptosis in human disease. A host cell containing a polynucleotide encoding CARD can be used to create transgenic animals.	Sequence	536 AA:	FT	Domain	7..98
C				FT		/note= "CARD domain"
C				FT	Modified-site	23..26
C				FT		/note= "Casein kinase II phosphorylation site"
C				FT	Modified-site	92..95
C				FT		/note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
O				FT		

XX	Bertin J;	Qy	459	-----LSATSSRELVDFRS---SSPAPPSSQQSISYKRAEDE	492
XX	WPI: 2001-167809/38.	Db	457	ggspkpkfaahqeqvlrphdagissgppbeker-rrikesf	499
DR	N-PSSDB; AA50387.				
XX	Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10, CARD-11, useful targets for therapy, as immunogens, and in screening and detection assays -				
PT	Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10, CARD-11, useful targets for therapy, as immunogens, and in screening and detection assays -	RESULT	8		
PT	Claim 9; Fig 5A-5B; 145pp; English.	ID	AB95617	standard; Protein; 366 AA.	
XX	The present sequence represents novel human caspase recruitment domain, CARD-9. The polynucleotide encoding this sequence was isolated from a human meckaryocyte cDNA library. Also described are novel human sequences for CARD-10 and CARD-11 (AAU01206, AAU01207) and rat CARD-9 (AAU01204). CARD-9, CARD-10 and CARD-11 interact with Bcl-10 which is thought to activate nuclear factor (NF)-kappa and apoptosis. The sequences of the invention can be used for treating a disorder associated with abnormal levels of apoptosis by modulating the expression or activity of CARD-9, CARD-10, or CARD-11. They can be used for the treatment of hyperproliferative disorders (e.g. cancer), autoimmune disorders (e.g. systemic lupus erythematosus), neurological disorders (e.g. Alzheimer's disease), inflammatory disorders (e.g. Crohn's disease), and viral infection (e.g. HIV). The CARD polypeptide, polynucleotide and antibody which selectively binds to CARD can be used in screening and detection assays (e.g. chromosomal mapping, tissue typing) predictive medicine (prognostic assays, monitoring clinical trials, and therapy (treatment and prophylaxis). The CARD polypeptide may be used to screen for drugs that bind to and/or modulate it. CARD sequences are potential targets for regulating inflammation, cancer, NF-kappa signalling, stress related response and apoptosis in human disease. A host cell containing a polynucleotide encoding CARD can be used to create transgenic animals.	PS	AB95617;		
XX	Sequence 536 AA;	AC			
XX	Query Match 9.6%; Score 493; DB 22; Length 536;	XX			
Matches 155; Conservative 92; Mismatches 20;	Pred. No. 9.9e-12;	XX			
Db	6 rddecvniedfrvtltvipsritpylrckvlnpddeeqvlsdpnlvirkkrkgvll	14;			
Qy	15 DEETLWMMEMSHRHIVRVCIPSRLTPYLRAKVLCQLDDEEVHSRPLTSNAMRAGHLL	74			
Db	66 dilqrteghkgvafiesleiytpqlykvgkeparvsfmidasqsgtq1lmtevmk	125			
Qy	121 TSKLTCLAGIGSLOELLNQKGKEVLLRCQQLQEHLGAETRAEGHLQLEADHSHRM	180			
Db	126 1qkkvqdltaillss-kddflkilrkvdslirkhgervq-----rl	164			
Qy	181 KREVSAAHFEYVLRKDEMILSLHYNSNAOKELAASRCRSLQEEYLQKELQRANMVS	240			
Db	165 keecegsrelkrckeeyndamlrahqseekgaaalmrnidqlieqdqkhslmiae--d	222			
Qy	241 SCELELQEQ-SURTASDQESDEELNRLKEE---NEKLRSLTFSLAEK-----DI	286			
Db	223 ckverkhtlkirhameqrpsqellwlgkallqarjeleasyqegkldrsspyiqv	282			
Qy	287 LEQSLDEARGSRQELVERIHSURERAVAAEQRQWEEKTQTLQFQKSIMACOLYREK	346			
Db	283 1eedwqalrdhrqeqgantifsrkdrggeerrlrcmekmfqclairkdsmykdr	342			
Qy	347 VNALQAOYCELOKERDOAYSADSQAETISSOLVKEKDSLROVFELTDQVCLRTOLRQL	406			
Db	343 ieail1qmeevaiercqatareelhqahqglqekdalrqrvelekaedlq1qvfqc	402			
Qy	407 QAEPPGVLKQOEARTRCPKPRERKQLYRMHAI---CPRDDSDSCSYSSTESOLSD----	458			
Db	403 eaq---llavegrirr---qpletivssldedsprrsqslsqdledtqisdkgcia	456			
Qy	Sequence 366 AA;	SQ			
Db	Query Match 7.7%; Score 396.5; DB 22; Length 366;	SQ			

Best Local Similarity 30.7%; Pred. No. 5e-24;	DR N-PSDB; AAI58090.
Matches 116; Conservative 65; Mismatches 146; Indels 51; Gaps 8;	XX
	PT Novel nucleic acids and polypeptides, useful for treating disorders
	PT such as central nervous system injuries -
Qy 15 DEETLNMMESRHRHTRVRCPSRLPILRQAKVQLQDEEEFLHSPRLTNSMARGHLL 74	XX
Db 6 nndeccavilegfrvtvtsivqsfsptrpylrgkvlnpddeeqvslapnvlirkvvgvll 65	XX
Qy 75 DLIKTRGKNGAIAFLESLKFRMPDVYTLVIGIQLPD----ME 120	XX
Db 66 dlqrighkgiyafvlesllyppqlkkykeparvfmildasgesqltqlmtevmk 125	CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.
Qy 121 TSKLTECLAGATGSLOPELNQDKGQKEVLLRRCQQLEHUGLAETRAEGLHQLEADHSRM 180	CC
Db 126 1qkkvqdltails-kddflkevkdslrlhgvqv-----rl 164	CC
Qy 181 KREVAISAFHEVFLRKDMLSLHYNALQENELAASRCRSQEEYLKLQELQRANMVS 240	CC
Db 165 keeceaadsrelrkckeenylamrnlahqseekgaa.lmrnrdlqleidqlkhslmkae--d 222	CC
Qy 241 SGELEQEQ-SLRTASDQESEGGDEELNLKEE---NEKLRSUTFSLAEK-----DI 286	CC
Db 223 dckverhtlkrrhamecrpsjellwlgallqarvqeasvqekidrspspiqv 282	CC
Qy 287 LEOSLDIARGSRQELVERIHSRERAVAAREQWEEKETOLLOFQKSMAQOLYREK 346	CC
Db 283 leedwqalrdhqeqantstirkdr1rqgeartrrlrcemeekemfelqclal;kdskmvkd 342	CC
Qy 347 VNLAQAOVCCELOCKERDOA 364	Sequence 350 AA;
Db 343 iealllgmeevaidqs 360	XX
	Note: The sequence data for this patent did not form part of the printed specification.
	Query Match Score 309; DB 22; Length 350;
	Best Local Similarity 32.2%; Pred. No. 8e-17;
	Matches 112; Conservative 50; Mismatches 148; Indels 38; Gaps 12;
Qy 668 LEAKVATSG-DSFYIRYVNLAMEGRAKE-LQYHCNEVLYHVTDTMFQCGCWMHAHYNSYT 725	XX
Db 13 leawakpqgaeplyiranitlperadphalevkageilrvdsaykrrewictrvdpt 72	XX
Qy 726 MKDTAAHTGTTIPHSRAOOQLTALIQDMTQCTVTRKPS- GGPOKLVRIVSMD---- 777	XX
Db 73 1rd-lrdgtvpnyqragql-----evqekcl---psstrhgrprsnkkraldqrlvr 122	XX
Qy 778 -KAKASPIRLSPDQGQDPSRNEGSTCFWAQSLTIVPYTULWPHPARPPVLLVPA 836	XX
Db 123 pkpvqgapqdpdqlllep-----c--aepersrpyslvrlpvlsairpvvlp 172	XX
Qy 837 VCKILSEKLCLLQG---FKRKCLAEYLQSQEYYEAWSQRGDIIQEVEYSGGRCWVTHRVAE 892	XX
Db 173 iaprlirnlldpssrlfdqvcpaeslsgelcpssqsgapqaqatpglisri-raiqe 231	XX
Qy 893 SLMEKNTNHALDYLQDLSVCTLHRMDIEPPIVTHVSNEKMAKKLKGJLQRLLSTSEEQLEA 952	XX
Db 232 svgkk--hcillelgaryrelvqneiyipqhvtevrgvlgpgrwdseilrql 289	XX
Qy 953 ARQEEGGLDRAFLQKLYSSLAPDWSLDLQGGLSCVRQTADEKKVVWTE 1000	XX
Db 290 crgseqywglpcswvqvpahewghaeelakvrgrilqeqarlvwve 337	XX
RESULT 9	RESULT 10
ID AAM38934 standard; Protein: 350 AA.	ID ABB64619 standard; Protein: 1445 AA.
XX AAM38934;	XX
XX 22-OCT-2001 (first entry)	XX
XX DE Human Polypeptide SEQ ID NO 2079.	XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; Alzheimropic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.	XX
XX Homo sapiens.	XX
XX WO200153312-A1.	XX
XX 26-JUL-2001.	XX
XX 26-DEC-2000; 2000WO-US342651.	XX
PF PR 21-JAN-2000; 2000US-0488725.	XX
PR 25-APR-2000; 2000US-0552317.	XX
PR 09-JUL-2000; 2000US-0598042.	XX
PR 19-JUL-2000; 2000US-0602312.	XX
PR 03-AUG-2000; 2000US-0653450.	XX
PR 14-SEP-2000; 2000US-0662191.	XX
PR 19-OCT-2000; 2000US-0693036.	XX
PR 29-NOV-2000; 2000US-0727344.	XX
PA (HYSEQ) HYSEQ INC.	XX
PA Tang YR, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	XX
PA Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Drmanac RT;	XX
PA Zhao QA, Zhou P, Goodrich R,	XX
XX (WPT.; 2001-442253/47.	XX

Db	263	linckgaaspelkeelrdkn-----esitmrlhetlkqcgdklktksqmdrk	312	AAG65799 standard; protein: 2101 AA.
Qy	64	TNSAMRAGHLILLLKTRKGIAF-TESLKEFHNPDVYTLVGLQDFVDFNFSGIMETS	122	ID AAG65799 XX AC AAG65799;
Db	313	inglese-----engdlsfkrefasshlqlgdaineltehskattgewlekqa	360	XX DT 30-JAN-2002 (first entry)
Qy	123	KUTECLAGAIGSLQEELNQEKGKEVILRLRCCQLQPHGLAETRAEGLHOLEADHSRMKR	182	XX DE Nuclear mitotic apparatus protein (NUMA) sequence.
Db	361	qlekeisaa---iqdktkcllee-kneillqkgisqleehls-----qldnpqcekg	406	XX KW Nuclear Mitotic Apparatus protein; malignant; cell growth; human; tumor progression; structural protein.
Qy	183	EYSAHFHEVLRKDEMISLS-----LHNSALQEKEELAASCRSRLQEEELYL-	228	XX OS Homo sapiens.
Db	407	evlgdvigletlkqeaatlaanntqlqarvemletergqgeakillaerghfeekqolss	466	XX PN US6287790-B1.
Qy	229	-----LKOBLQRANNMVSCELELQEQSRT-----ASDQECSDEEINRLK	268	XX PD 11-SEP-2001.
Db	467	11tdlqssisnlsqakeelesqasqahgaritaqvasitselftnatiqq-qdqelaglk	525	XX PF 30-NOV-1999; 99US-0452294.
Qy	269	EE-----NEKLRSLTFSLAEKDILEQSDLEARGSRQELVERIH	306	XX PR 30-NOV-1998;
Db	526	qjakekaqjaqtllqqeqasglirngveqissllqk---eqqile-re-aekqeatrqdn	581	XX DR WPI: 2001-610603/70.
Qy	307	SLRERAYAAERREQYWEKEETLQLQFOKSMAQCYREKYNALQVQCELOKERQDAYS	366	XX PS Disclosure; Fig 15; 38pp; English.
Db	582	a;qjlataae-ereasretdalkaqe-----alekekaakae----qqqlqvane	629	XX PA (REGC) UNIV CALIFORNIA.
Qy	367	ARDSAQREIISQSLIVEKRSLSLRQYFEL-----TDOVCELTQRLQNLQEPFG	412	XX XX
Db	630	ardsaqstvtqsgrekaelsrveelqacvetarqeqheadaqeqaelqirseq----	684	CC Distinguishing growth arrested malignant cells from proliferating non-malignant cells, comprises staining the nuclear mitotic apparatus protein in intact mammalian cells and determining the 3-dimensional pattern of the labeled protein -
Qy	413	VHQEARTRPCEPREKORLVRHAICPRDDSSCSLYSSTESTSULSDLSATSSRELVDTSFR	472	XX DR
Db	685	--qkaterkerryaqekq-----qeqjalkesklskrgsleekr	723	XX PS
Qy	473	SSSPAPPSSQSLSYKRVAAEDFGEEPWFSFSSCLIEPEGDPGALPFAKACDPHLYELLDTD	532	XX CC
Db	724	724 radialeeqqrctiseikae-----trs1vbgkrkrkeleergrkglearllqgle	776	CC CC
Qy	533	LPOLESSLQPVSPGRFLDVSEGVIMRRPARBLSQVTMMLAFQGDALFQISVIGNITG	592	CC CC
Db	777	777 anaete-----virrelaemaaqht-----aeseceql-----806	CC CC	
Qy	593	TFIHRVTPTGSAADQMALRPGTQIVMDYEASBPLFKAVLEDITLEEAVGLLERVDGFCCIL	652	CC CC
Db	807	-vteavaardqye----dsqeqeabaqymfqeq-----lm 836	CC CC	
Qy	653	SVKVNTDGYKRLIQDLEAKVATPSYVIRVNLAMEGRAKGEQVHCNEVLLVTDIMFQG	712	CC CC
Db	837	tiveeekarqeqeakkva--giesselsqisrqgnklael--han-----880	XX CC	
Qy	713	CGCWHAHRVNSTKMDTAAGHTPNYSRAOOOLIALIDMTQCTYTRKPSSGGPKLV-	771	CC Sequence 2101 AA;
Db	881	881 ---laraiqgyqkevraqkladd1stlqemnaatskevarletivrk--ageqgetas	934	CC SQ
Qy	772	RIVSMDKAKASPLURLSFDRGQDLSRME---GSSTCWAESCLTVPTLYWPHRPARPR	828	Query Match 5.1%; Score 262; DB 22; Length 2101; Best Local Similarity 19.9%; Predict. No. 1.1e-11;
Db	935	relvkepara-----qdcpewleeqqrqfc-----stqa	965	Mismatches 208; Conservative 163; Indels 296; Gaps 39;
Qy	829	PVLLVPRAVGKILUSEKUKLQQFKKCLAEYLQOEYEWASQGDIIQEGEVSGRCWVTR	888	Qy 4 LCRRDSALTADELTMWEMMESHRRIVRCIPSRLTPYLQAKVLQCLDDEEVLHSRPL 63
Db	966	966 alqamereaeqmgne----leriraalmesqgqgeerqgeeraritqerg-----r 1015	Db 263 linnekqaaasp1lepkelieelrkn-----estlmurheli1kqcdlkteksgmrd 312	
Qy	889	HAYESLMEEKNTHALDVOOLDSTVCLTHRMDFIPF--VIHVSVNEKMAK-----KLK-----936	Qy 64 TNSAMRAGHLILLLKTRKGIAF-LESLKRHNPDVYTLTGLQPDYDFSNFSGLMETS 122	
Db	1016	1016 agadlaalekaarcademrlnqalnegrvefatqeaalahaaltekegdqelakrlgleaa	1075	Db 313 inqlse-----engdlskfklerefashlqlqdainelteehskatgewlekqa 360
Qy	937	--KGLQRJGTSEFOLLEAARQEE	957	Qy 123 KUTECCLAGAIGSQEELNQEKGOKEVILRRCOOLQEHNLGLAETRACQHOLEADHSRMKR 182
Db	1076	1076 qikeleerqtvqkqiegaklaje 1098	Db 361 qikeleasa---1qdktkcllee-kneilqgklsqleehns-----qldnpqcekg	406
Qy	229	-----LKQELQRANMVSSCELQEQSLRT-----ASDQESGDEELNRK	268	Db 407 evlgdvigletikqeaaatilaantqiqarvemletergqgeakllaerghfeekqgqis 466

RESULT	14			
AAV49936				
ID	AAV49936	standard; protein;	2101 AA.	
XX	AAV49936;			
AC				
XX				
DT	01-FEB-2000	(first entry)		
XX				
DE	Human	NUMA	protein #1.	
XX				
KW	KW	KW	KW	DNA-activated protein kinase; catalytic subunit; PARP; detection;
KW	KW	KW	KW	nuclear NAD+ ADP-ribosyltransferase; autoantigen; immune response;
KW	KW	KW	KW	autoimmune disease; cancer; type I diabetes mellitus; thyroiditis;
KW	KW	KW	KW	myasthenia gravis; primary biliary cirrhosis; rheumatoid arthritis;
KW	KW	KW	KW	systemic lupus erythematosus; polymyositis; dermatomyositis;
KW	KW	KW	KW	Sjogren's syndrome; scleroderma; graft-vs-host disease.
XX				

OS Homo sapiens .
XX WO9953757-A1 .
PN XX
PD XX
28-OCT-1999 .
PF XX
22-APR-1999 ; 99WO-US05774 .
PR XX
22-APR-1998 ; 98US-0082643 .
PA XX
PA (MERCK & CO INC .

PA (UYJO) UNIV JOHNS HOPKINS .
PA (MERI) MERCK FROSST CANADA INC .
XX
PI Thornberry N , Rosen A , Casciola-Rosen L , Andrade FA , Nicholson D ;
PI Roy S ;
XX DR
XX WPI; 2000-013162/01 .
XX Autoantigenic fragments useful for diagnosis, treatment of autoimmune diseases and cancer -
XX PT PT
PS PS Claim 4 : Fig 9 ; 88pp ; English .
XX The present invention describes autoantigenic fragments (I) produced by the action of a lymphocyte granule enzyme (II) on isolated cells containing an autoantigen. (I) is useful for prophylactic and therapeutic treatment of an autoimmune disease (AI) such as type I diabetes mellitus (I), thyroiditis, myasthenia gravis, primary biliary cirrhosis, systemic lupus erythematosus, rheumatoid arthritis, polymyositis, dermatomyositis, Sjogren's syndrome, scleroderma and graft-vs host disease. (I) produced by the action of (II) on cells isolated from target tissue is administered to a patient for tolerising to the presence of (I). As a therapeutic treatment, isolated (I) associated with AI condition is contacted with serum of the patient containing autoantibodies (AAB) against (I), under condition to allow the in vivo binding of Ab to (I) and a portion of the Ab is removed from the serum of the patient. (I) derived from malignant cell is administered to a patient for stimulating an immune response against malignant cells to treat cancer. (II) is also useful in assays for assessing the presence or absence of an AI condition in a patient. Such assays involve detecting AAB in a sample by using (I) or detecting (I) in a sample by using antibodies that specifically binds to a cryptic epitope of (I). The present sequence represents the human NumA protein, which is a specifically claimed autoantigen from the present invention .
Sequence 2101 AA ;

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 14061; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent
 CC useful in detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL0511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABL57737-ABL72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1916 AA;

Query Match 5.0%; Score 258; DB 22; Length 1916;
 Best Local Similarity 22.1%; Pred. No. 2e 11; Mismatches 97; Indels 164; Gaps 27;
 Matches 133; Conservative 97; Gaps 27;

Qy 473 SSSPAPPSSOOSLYKRAVAEDFGEPEPWFSFSCLEIPCEGDPALPGAKAGDPHDLY- 532
 Db 724 raadaleegqrcciseikae----trsiveghkrkeleeragrkglearllqge 776
 Qy 533 I P Q L E S S L O P Y S P G R L D V S E S G V L M R R P A R I L S Q V T M L A F Q G A L L E Q I S V I G G N L T G 592
 Db 777 ahqaete----virlrealaemaqaqt---aseedcl---- 806
 Qy 593 I F I H R V T P G S A D O M A L R P G T Q I V V W D Y A E S P E L K A V L E D T T L E B A V G L I R R V D G F C L L 652
 Db 807 ---wkeavaawrdyee----dsqeqeaqrfqeq-----1m 836
 Qy 653 S V K V N T D G Y K R L Q D L E A K Y A T S G D S F Y I R V N L A M E G R A K G E L Q V H C N E V I L H V T D T M F Q G 712
 Db 837 tkeecækarcælqeakekra--glehselqisrqqnkael--han---- 880
 Qy 713 C G C W H A H R V N S Y T M D T A A H T I P Y S R A Q C Q L I A I Q D M T Q C V T R K P S G G P Q K L V - 771
 Db 881 ---laralqpvkevraqiajaddlqekmaatskevarletivrk -sgeqetas 934
 Qy 772 R I V S M D K A K A S P L R L S F D R G Q L D P S R M E -- G S S T C F W A E S C L T V P T L W P H R P A R P R 828
 Db 935 relvkpara-----gqrqpwleeggqrqfc-----stqa 965
 Qy 829 E V L V Y P R A V G I L T K I L S E K U C I L L G F K R C L A E Y I S Q Q E V E A W S R G D I I Q E G E V S G G R C W V T R 888
 Db 966 alqamereaaemgne----lerlraalmessggqgeergggerevarltgerg----r 1015
 Qy 889 H A V E S T M E K N H T A L I D Y Q D S V C T U H R M D I F P I - V I H V S Y N E K M A K - - - - K U R K ----- 936
 Db 1016 aquadialekaraeIemrlqlnahnevratqalahaitekegkdgelakirglesa 1075
 Qy 937 -K G L Q R L G T S E Q O L L E A R O B E 957
 Db 1076 qikeelrqrqvqkq. keglakirke 1098

RESULT 15
 ID ABB62423 standard; Protein; 1916 AA.
 XX ABB62423;
 XX 26-MAR-2002 (first entry)

DE *Drosophila melanogaster* polypeptide SEQ ID NO 14061.
 XX
 KW *Drosophila*; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX OS *Drosophila melanogaster*.
 XX PN WO200171042-A2.
 XX 27-SEP-2001.
 XX PR 23-MAR-2001; 2001WO-US092231.
 XX PR 23-MAR-2000; 2000US-191633P.
 PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI: 2001-656860/75.
 DR N-PSDB: ABL06526.

Result No.	Score	Query Match	Length	DB ID	Description
1	262	5.1	2101	1 US-08-466-390-4	Sequence 4, Appli
2	262	5.1	2101	1 US-08-470-950-4	Sequence 4, Appli
3	262	5.1	2101	1 US-08-467-781-4	Sequence 4, Appli
4	262	5.1	2101	1 US-08-193-487-4	Sequence 4, Appli
5	262	5.1	2101	2 US-08-483-924-4	Sequence 4, Appli
6	262	5.1	2101	4 US-09-294-294-1	Sequence 4, Appli
7	262	5.1	2101	5 PCT-US93-06160-4	Sequence 4, Appli
8	238	4.6	3248	1 US-08-333-700-1	Sequence 1, Appli
9	238	4.6	3248	5 PCT-US95-16236-1	Sequence 6, Appli
10	237	4.6	2482	1 US-08-318-254-6	Sequence 6, Appli
11	216.5	4.2	1388	2 US-08-655-576-1	Sequence 1, Appli
12	210.5	4.1	976	4 US-09-104-324-4	Sequence 6, Appli
13	205	4.0	816	2 US-08-533-3068A-6	Sequence 6, Appli
14	205	4.0	816	2 US-08-732-923A-6	Sequence 6, Appli
15	205	4.0	885	2 US-08-533-306A-4	Sequence 4, Appli
16	204.5	4.0	885	2 US-08-732-922A-4	Sequence 4, Appli
17	204.5	4.0	1354	3 US-08-655-871-2	Sequence 2, Appli
18	200.5	3.9	1388	2 US-08-655-576-4	Sequence 4, Appli
19	197	3.8	576	2 US-08-533-3068A-2	Sequence 2, Appli
20	197	3.8	576	2 US-08-732-923A-2	Sequence 2, Appli
21	197	3.8	1886	4 US-08-938-105-3	Sequence 3, Appli
22	195	3.8	1090	4 US-09-005-199B-5	Sequence 5, Appli
23	194	3.8	1388	4 US-09-532-191-2	Sequence 2, Appli
24	192.5	3.7	756	4 US-09-005-199B-9	Sequence 4, Appli
25	190	3.7	914	4 US-09-005-199B-9	Sequence 4, Appli
26	189	3.7	1898	1 US-08-056-200-94	Sequence 94, Appli
27	189	3.7	1898	2 US-08-800-644-94	Sequence 94, Appli

Qy 64 TNSMRAGHLLDKTRKGNGAAIAF-LESLKFHNPDYTYLVLQPDVFDFNSFGSLMETS 122
 Db 313 INQSLSE-----ENGSFKLREFASHLQQLDANELTBHSRATQEWNLKQA 360
 Qy 123 KITECLAGAIGSLOQEELNQEQKEYLRLRRCQQLQPHGLAETRAEGLHQLEADHSRMKR 182
 Db 361 QLEKELSAA--LQDKKCLEE-KNEILQGKLSQLEELHS-----QLQDNPQPKG 4 06
 Qy 183 EVSAHFRREVLRKDEMULSLS-----LHSNALQBKEELASCRSRLQEEELYL-- 228
 Db 407 EVLGDVQLETIKQEATLAINTLQARVEMLETERGQEAQKLLAERGHPEEERQQLSS 4 66
 Qy 229 -----LKOELQRANMVSSCCELEPEQSLRT-----ASDQEGSDEELNRK 268
 Db 467 LTDLOSSISNLSQAKELEQNSQAHGARUJTAQAVSUTSETLTNTACQOO-DQEELAGLK 5 25
 Qy 269 EE-----NEKLRSLSUTSLAEKDILEQSLDEBARGSRQELVERIH 306
 Db 526 QQAKEKAQQLACTLQQEQQASQGLRIQEOLSSLQK--EQQLEKV-AEKQEAETRDH 5 81
 Qy 307 SIRERAVAAREQWEEKETTOLQFQSKMACOLYREKVNAQOVCLEQKERDQAYS 366
 Db 582 A-QQLATAAE-FREASLERAAKLE-----ALEKERAAKLET---LQOQLOVANE 6 29
 Qy 367 ARDSAQREISQSLVEKDSLRRQEVFL-----TDQVCEBLTQLRQLOREPPG 4 12
 Db 630 ARDSAQTSVTQDQREKAELSKEELQACVETARQEOHEAQVAQABLELQRLSEQ---- 684
 Qy 413 VLKQEARTRPQPREFKORLVRMHAICPRDDSDCSLYSTSSESQQLSDISATSRELVDSSFR 4 72
 Db 685 --QKATEKERVAAQEDQL-----QEQLQALKESLKTVKGSLEEKR 7 23
 Qy 473 SSSPAPPSSQSLSYKRAEDEGEPWSFSSCLELIPEDPGALPAGAKAGDPHDYELDTAD 5 32
 Db 724 RADALEBQORCISELKAE-----TRSILYEQHKERKELEERGRKGLEARLQLGE 7 76
 Qy 533 LEPOLESSLQPVSPGRLDVSESSVLMRMRPARIISQYTMIAFGODALLEQTSVGGNLTG 5 92
 Db 777 AHOAETE-----VLRRELAFAMAQHT-----AESECEQL-----806
 Qy 593 IPTHRYTPGSSADOMALRPGTQIVMDYEASPLPKAVLEDTTLEAVGLLRRVDGFCCL 652
 Db 807 -----WEVAAMRDGYE--DSQQEEAQFCAMFQDQ-----LM 8 36
 Qy 653 SYVKVNTDGYKRLQDLEAKVATSGDSFYIRVNLMAGRAXKQEVHNEVLTQDTMFQG 7 12
 Db 837 TKEECEKARQLEKEAII-----GTEBSHSELQISROONKLAEL-HAN-----880
 Qy 713 CGCWHAHRVNSTMKDTAAHTIPNYSRAQQQLIAIJDMTQCTVTRKPSSGGPKLV- 7 71
 Db 881 -----LARALQOQYEKEVRAQKADDLSTLQEMMAATSKEARLETLYRK--AGEQETAS 9 34
 Qy 772 RIVSMDIKASLRLSLPDRQDLPMSME--GSSTCWAESCLTLVPTLYWPHRPAPR 8 28
 Db 935 RELVKEPARA-----GDRQPEWLEEQGRQFC-----STQA 9 65
 Qy 829 PVLVPRAVGKILSEKULCLLOGFKKCLAEYLQEGEVGGRCWVTR 8 88
 Db 966 ALQAMERAEQGN-----LEURAAIMESQQQQERGQDPERARLTQDFR-----R 1015
 Qy 889 HAVESLMEKNTHALDVOOLDSTCLHRMDIFFI -VHVSYNEKMAK-----KLK----- 9 36
 Db 1016 AQADALEKAARAELENRLQNLNAINEQRVFATQEAIALAHTEKEGRDQELAKRGLEAA 1075
 Qy 937 -KGQLRGTSSEBOLBAAROE 957
 Db 1076 QIKEELEROTVKQLKEQQLAKKE 1098
 Result 2
 US-08-470-930-4
 Sequence 4, Application US/08/70950
 Patent No. 5698439

GENERAL INFORMATION:
 APPLICANT: TOUKATY, GARY
 APPLICANT: LIDARD, GRAHAM P
 TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TESTA, HURWITZ & THIBEAULT
 STREET: 125 HIGH STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,950
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 516
 ATTORNEY/AGENT INFORMATION:
 NAME: PITCHER SQ, EDMUND R
 REGISTRATION NUMBER: 27,829
 REFERENCE/DOCKET NUMBER: MTP-013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2101 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-470-950-4

Query Match Score 262; DB 1; Length 2101;
 Best Local Similarity 19.9%; Pred. No. 5.7e-13;
 Matches 208; Conservative 163; Mismatches 376; Indels 296; Gaps 39;

Qy 4 LCRDSDSATLDEETLWMMBSSHRRHIVRCIPSRLTPYLEQAKVLQLDDEBVHLSPRL 63
 Db 263 LLNEKQAAASPLEPKELFLRDKN-----ESUTMLRHETLKCQCDLKTEKSQMDRK 312
 Qy 64 TNSMRAGHLLDKTRKGNGAAIAF-LESLKRFHNPDVYTLYTGLQPDVDFNSFGIMETS 1.22
 Db 313 INQLE-----ENGDLFSKLFREFASHLQLQDNLNEIEHSKATQEWLEKQA 360
 Qy 123 KTECLAGAIGSLQEELNQEQKEVLLRRCQQLQEHGLAETRAEGLHOLEADHSRMKR 182
 Db 361 QLEKELSA--LQDKKCLEE-KNEILQGKLSQLEELHS-----QI-QDNPQKEKG 406
 Qy 183 EVAHFEVLRKDEMILS-----LHYSNALQEKELAASCRSRLQEEELYL-- 228
 Db 407 EVLGDVQLETLKQEAATLAANTQLOQARVMLETERGQEQPKLAAERGHFEEEQQLSS 466
 Qy 229 -----LQELQRANMVSCELELQEQSLR-----ASDQEGSDEELNRLK 268
 Db 467 LITDQSSISNLSQAKEELEQASQAHRLTQASVATLSLTLNATIQQ-DQELAGLK 525
 Qy 269 EP-----NEKLRSLTFSLAEKDILEQSLDEARSQELVERIH 306
 Db 526 QKAEKOQALQACTLQOOQEQASQGLRHOVQUSSSLQK--EQQLEKV-AEKQEAETRDH 581
 Qy 307 SLRERAVAFAERQEQYWEKEKTLQFQSKMACOLYREKYNALQAVQCELQERDOAYS 366
 Qy 367 ARDSAQETISOSLVERDSLRQFVFL-----TDQVCELRQTLQQLQEPG 4 12
 Db 630 ARDSAQTSVTAQREKAELSKEELQVQVAALELQLRSEQ---- 684

413 VLKQEARTREPCPREKQLRVLRMHAICPRDDCSLVSSESOLLSLSDSATSSRELVDTSFR 472
 685 ---QKATEKERAQEKEQDOL----- : | : | : | : | : | : | : | : |
 473 SSSPAPPSQSOLSYKRAVEDFGEPPWSFSSCLEIPEGDPCALPGAKGDPHDYELLDTAD 532
 724 RAADALEDEQRCITSELKAE-----TRSLVEQHKRERKELEERGRKGLEARILQLGE 776
 533 LPQLESSILOPSPGRLLDSESGYLMRRPARRILSQTIVMLAFOGLLBOISVGGNLTG 592
 777 AHOAETE-----VLRLAELAAMAQHT-----AESECEOL----- 806
 593 IFIHRVTPGSAADQMAILRPGTQIYVMVDEASEPLFKAVLEDTTLFEEAVGLRLRYDGFCCCL 652
 807 -----VKEVAANWRDGYE---DSQEEAEQYGMFQE-----LM 836
 653 SYKVNTDGYKRILQDFKAATSGDSFYTRVNIAAMEGRAKGELQVHNCNEVLHVTDTMFQG 712
 837 TLKEECCKARQEAEKVA -GIESSELQISRQNKLAEEL -HAN----- 880
 713 CGCWHAHRVNSYTMKDPAAHGTIPNYSSRAQQQLIALQDMTOCCTVTRKPSSGGPQKLV- 771
 881 ----LARALQOYQEKEYRAQKLADDLSTLQEKAATSKEVARJETLVKR--AGEEQETAS 934
 772 RIVSMDAKASPLRLSPFDRGOLDPSRME---GSSTCFWAESCLTLVPTLYWPHRPARPR 828
 935 RELVKEPARA-----GDROPEWLEQQGRC-----STQA 965
 829 PVLVPRAVGKILSEKCLLQGFPKKCIAEYLSSQEYEAWSQRGDLIQEDEVSGGRCWVTR 888
 966 ALQAMEEAEQNGNE----LERLRAALMSEQGOCQERQOEREVARLTQERG-----R 1015
 889 HAVESTMENTHALLDQDSDYCLTHMDIFPI -VIHVSNEKMAK-----KELK----- 936
 1016 AQADLALEKAARAAELENURLQNALNEQRVEFATLQEAHALTEREGKQDQELAKLRGLEAA 1075
 937 -KGQLQFGTSEBOLLAEARQEE 957
 1076 QIKEELEROTVQKLKEQLAKKE 1098

RESULT 3
-08-467-781-4

Sequence 4 , Application US/08467781
 Patent No. 5780596
 GENERAL INFORMATION:
 APPLICANT: TOUTRATLY, GARY P
 APPLICANT: LUDGARD, GRAHAM P
 TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSSEE: TESTA, HURWITZ & THIBEAULT
 STREET: 125 HIGH STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02110

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,781
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: PITCHER ESO, EDMUND R
 REGISTRATION NUMBER: 27,829
 REFERENCE DOCKET NUMBER: MTP-013
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2101 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-467-781-4

	Query Match	Score	DB 1:	Length	2101;
	Best Local Similarity	19.9%	Pred. No.	5.7e-13;	
	Matches 208;	Mismatches	376;	Indels	296;
	Conservative			Gaps	
Qy	4 LCRDSDAHLDDLLKTRKGNGAIAF-LESLKHEHNPDVYTYLGLQPDVFDSNFSGLMETS	63			
Db	263 LLNEKQAAASPLEPKELLELRDKN-----				
Qy	64 TNSAMRAGHLLDKTTRKGNGAIAF-LESLKHEHNPDVYTYLGLQPDVFDSNFSGLMETS	122			
Db	313 INQSL-----ENGDLSFKEFLREFASHLQQQLDAINELTEBHSKATQEWLEKQA				
Qy	123 KLTCELAGAIGSLSQEELNQEKGQKEVLLRRCQQLQFHLGLAETRAGLHQLEADHSRMKR	182			
Db	361 OLEKELSAA--LDQKCKLCE-KNEILQKGKISQLEEBLHS-----QLQDNPPQEKG	406			
Qy	183 EVSAHFHEVLRKDEMELIS-----LYHNALQNEKELAASRCSRRLSQEEELYL-	228			
Db	407 EVLGDVLOLETIKQEATLAANNTQLQARVEMLETERGQOZAKLLAERGHFEEEKQQLSS	466			
Qy	229 -----LKOELQRANMVSSCELELQEOSLRT-----ASDQESEGDEELNRLK	268			
Db	467 LITDLOSSINSQAKEELQASQAHGARLAQAVASLTSELTTLNTAQOO-DQELAGLK	525			
Qy	269 EE-----NEKLRSLTFSLAEKDILQFDLSDEARGSRQELVERIH	306			
Db	526 QAKEOQAQLAQTLOQQEQAQSGLRHRQEVSQSSLRQK-----EQOLKKEV-AEKKQETRQDH	581			
Qy	307 SLRERAVAAREQWEEKETLLOFQKSXMACOLYREKVNALQOVCZELOKERDQAYS	366			
Db	582 A-QQLATAAE-EREAStRERAAALKQE-----ALEKEKAAKLET-----LQQQLOVANE	629			
Qy	367 ARDSAQEISQSLVKEDSLRRQYFEL-----TDQVCELRTRQQLRQLQAEPPG	412			
Db	630 ARDSAQTSTVQAQREKAELSKVVEELQACVETARQEHEAQQAQVALLQLRSEQ-----	684			
Qy	413 VLKQEARTREPCKPREKORLYVRMHAICPRDDSDCSLVSSTESQQLSDLSATSSRELDSFR	472			
Db	685 --QKATEKERVQAQEDQL-----QEQQJALKSLSKTVKGSLEBEKR	723			
Qy	473 SSSPAPPSQSLKYKRYADEFGEPEWPSFSSCLFIEPGDPGALPGAKGDPHDYELLDTA	532			
Db	724 RAADALEEQRCISELKAE-----TRSYLVEHQKRKELEERAGRKGLEARLQLGE	776			
Qy	533 LPQLESSLPQYSPGRDVSSESSVLMRRRPARRILSQTVMFAQDALLQIVSGVNLTG	592			
Db	777 AHOAETE-----VLRRELAFAMAQHT-----AESEECQL-----	806			
Qy	593 IFTHRVTPGSAADOMALRPGTQIVMDYEASEPLEFKAVLIEDTTLFEEAVGLLRRVDGFCC	652			
Db	807 -----VKEVAAWRDGVE-----DSQQEBAQJGAMF95Q-----LM	836			
Qy	653 SVKVNTDGYKRLQLDEAKVATSGDSFYIYRNLANMEGRAKGELQVHCNEYLVHVTDTMFQG	712			
Db	837 TLKKECEKARQFLQEAKEVWA-GIESHSFLQISRQNKLAEL--HAN-----	880			
Qy	713 CGCWHAHRVNSYTMKDTAAHTGIPNNSRAQOOLAIQDMTQCTVTRKPSGGPKLV-	771			
Db	881 ---LARLQQYQEKEYVRAQKLADDLSTLQKEKMAATSKEVARLETLYRK--AGEQOETAS	934			
Qy	772 RIVSMKDAKASPLRSLFDRGQLDPSRM--GSSTCFWASCLTYLWPHRPARP	828			

Db 935 RELVKEPARA-----GDRQPEWLEQQGRQFC-----STQA 965
 Qy 829 PVILLEPRAVGKILSERKUULLQGFKCLAEYLSCOEYDWSQGDIIQEGEVSGRCMVTR 888
 Db 966 ALQAMERAEAOGN-----LERRAALMESQOQEQERGQPEREVALTUQES-----R 1015
 Qy 889 HAYEVSEMEKNTHALLDYOLDSCTLHRMDIFP-----VTHVSNEKMAR-----KLK-----936
 Db 1016 AQADLALEKAARRELEMRLQNALNEQRVEFATQEAALAHALEREKGDOELAKLRGLDEAA 1075
 Qy 937 - KGLQRIGTSEPOLLEAARQEE 957
 Db 1076 QIKELEELRQTVQLKEOLAKE 1098

RESULT 4

US-08-195-487-4

; Sequence 4 , Application US/08195487

; Patent No. 5783403

; GENERAL INFORMATION:

; APPLICANT: TOUKATLY, GARY

; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE

; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TESTA, HUBWITZ & THIBEAULT

; STREET: 53 STATE STREET

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/195,487

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/901,701

; FILING DATE:

; ATTORNEY/AGENT INFORMATION

; NAME: PITCHER ESO, EDMUND R

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: MTP-013

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/248-7100

; TELEFAX: 617/248-7100

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2101 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-195-487-4

Query Match 5.1%; Score 262; DB 1; Length 2101;
 Best Local Similarity 19.9%; Pred. No. 5.7e-13; Mismatches 376; Indels 296; Gaps 39;

Qy 4 LCRDSDSALTALDEETWMEMMESHRRHRRVRCPSRLTPYLAQAKVLQDDEBLHSPLR 63
 Db 263 LNEKOQASPLPEKELRUNK-----ESUTMRHLHETLKQCDLKTKEKSMDRK 312
 Qy 64 TNSAMRAGHLLDLKTRGNGIAF-LESIURHNPDVYTLYVQDPPDFDSNFSGLMETS 122
 Db 313 INQSL-----ENGDSLFSKIREFASHLQLQDNLNELLIEHSRATQEWLEKOA 360
 Qy 123 KLTCELAGAIGSLOQEELNQERGQKEVLLRRQCQLQEHGLAETRAEGLHOLEADHSRMKR 182

Db 361 QLEKELSSA---LQDKKCCLEE-KNEILQGKLSQLEEHHS-----QLDNPPQBKG 4 06
 Qy 183 EVSAHFHEVLRKDEMMLS-----LHYSNALQEKELAASCRSLSQLEELHHS-----LQDNPQBKG 228
 Db 407 EVLGDVLGLETLKQEAATLNAANTQOLQARVEMLETERQEQEAKLLAERGHFEKEQKQLSS 466
 Qy 229 -----LKQELQRANMVSSCELEQESLRT-----ASDQESEDEELNRLK 268
 Db 467 LTDLQSSISLNSLNUQSKELEQLEQASQAHARLTAQVASLTSSELTLNATQQ-DQELAGLK 525
 Qy 269 EE-----NEKRSLTSFLAERDILEQSLDERSGROELVERIH 306
 Db 526 QQAKEKQNQLAQTLQQQQEASQGLRHOEQLSSLKOK--EQQLKEV-AEQEAUTQDH 581
 Qy 307 SURERAYAERQREQYMEKEOTLQPKSKMAQLYREKVNAHQYCELOKERDOAYS 366
 Db 582 A-OQLATPAE-EРЕАСЛУРДАЛКЛЕ-----ALEKREAKLAI ---LQOQLQVANE 629
 Qy 367 ARDSAQRBISQSLVKEKDSLRLRQVFEL-----TDQVCCLRDTQLRQAEPPG 412
 Db 630 ARDSAQSIVTQREKAELSRVYELQACVETARQEHQEAQVAELQRLRSEQ----684
 Qy 413 VIKQEARIREPCPREKORLVRMHAICPRDDSDCSLVSYSTESOQLSDLSATSRELVDSSFR 472
 Db 685 ---QKATEKERYAQEKDQL-----QEQLQALKESLKVKIGSLEBEKKR 723
 Qy 473 SSSPAPPSQQLSYKRAEDFGEPWSESCLEPIEGPGALPGAKAGDPHDYELLDTD 532
 Db 724 RAADALEEQRCITSELKAE-----TRSLVQHKRKRKELBEBERAGRKGEARLQJGE 776
 Qy 533 LPQLESSQPVSPGRLDVSSEGVLMRRPARRLSQVMTMFAQDALLEQISVIGGNLTG 592
 Db 777 AHQAETE-----VLRLEAEMAQAH-----AEEC EQL-----806
 Qy 593 IFTHRVTPGSAADQMLRPGTQIVMDYEAESPPLFKAVLIEDTLEAVGLLRRVDFCCL 652
 Db 807 -----VKEVAAVRQDGVB-----DSQQEFAQYGMFQE-----LM 836
 Qy 653 SYKVNTDQYKRLQDIAKAVTSGDSPYIIRVYAMEFRAKGEQVHNELEVLTDMFQG 712
 Db 837 TLKEECEKARQELEQAEKVA-GTESHSELQISRQONKLAEL-HAN-----880
 Qy 713 CGCWHAHRVNSTYMKDTAAHGTIPNYSRAOOQOLIALQDMTOQCTYTRKPSSGGPKLV- 771
 Db 881 ---LARNLQVOKEVYPAQKLADDISTQEXMAATSKEVARLETIVRK--AGEQETAS 934
 Qy 772 RIYSMDKAKASPLRLSPDRGOLDPSRM-----GSSTCFWAESCLTLYVTPYLMPHRPAPR 828
 Db 935 RELVKEPARA-----GDROPWEEQOGQRFC-----STQA 965
 Qy 829 PVLLVPRAVGKILSEKULLQDFKKCLAEYLSQEEYBAWSQGDILIOGETYVGGRWVTR 888
 Db 966 ALQAMEREEAEQMGNE---LERAALMESQGQOQERGQERREVARTQERG-----R 1015
 Qy 889 HAVESLMEKNTHALLDYQDVSCTLHRMDIFFP-VIHVSNEKMAK-----KLK-----936
 Db 1016 AQDIALEKAARAEELRQLQNALNEQRFATLQEAHALTEKEGRDQELAKLRSLEAA 1075
 Qy 937 --KGLORGTSPEQLEARQEE 957
 Db 1076 QIKELEELRQTVQLKEOLAKE 1098

RESULT 5
 US-08-483-924-4
 ; Sequence 4 , Application US/08483924
 ; Patent No. 5883876
 ; GENERAL INFORMATION:
 ; APPLICANT: TOUKATLY, GARY
 ; APPLICANT: LIDGARD, GRAHAM P
 ; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 ; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

NUMBER OF SEQUENCES : 6
 CORRESPONDENCE ADDRESS :
 ADDRESSEE: TESTA, HURWITZ & THIBEAULT
 STREET: 125 HIGH STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/483, 924
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: PITCHER ESO, EDMUND R
 REGISTRATION NUMBER: 27,829
 REFERENCE/DOCKET NUMBER: MTP-013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2101 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-483-924-4

Query Match Score 5.18; Score 262; DB 2; Length 2101;
 Best Local Similarity 19.9%; Pred. No. 5.7e-13;
 Matches 208; Conservative 163; Mismatches 376; Indels 296; Gaps 39;

Qy 4 LCRDSLTADEETLWEMMESHRRIVRCICPSREITPYLRAKYVLCQLDDEEVLHSPLR 63
 Db 263 ILINEKQAASPLEPKELLELRDKN-----ESTLMRLHETLKQCDQDLTEKSQMDRK 312
 Qy 64 TNSMRAGHLIDLKLRKGNAIAF-LESFKHNPDVYTLYVGLQDPDVDFNSFSGLMETS 122
 Db 313 INQLE-----ENGDSFKLREFASHLQLQDALNELTEEHSAATQENWLEKQA 360
 Qy 123 KUTECLAGAISGLQEEENQGKEYLLRRCQOLQHGLAETRAEGLHOEADISMKR 182
 Db 361 QLEKELSAA---LQDKKCLE-KNEILOGKQSLEBHS-----QIQDNPPQEKG 406
 Qy 183 EYSAHPEVLRUKDEMISLS-----LHYSNALQKEKELAASRCRSLOQEELYL-- 228
 Db 407 EVLGDVYLQLETLIKQEATLAANNTQLQARVMLETGRQQAKLAERGHEEEKQLOSS 466
 Qy 229 -----LKOELQRANMVSSCELLQEQSLRT-----ASDQEGSDEELNRK 268
 Db 467 LITDQSSISNLSQAKEELQASOAHGARLTAVASHTSETLNTNATQQ-DQEVLAGLK 525
 Qy 269 EE-----NEKLRSQSLDPLARGQELVERIH 306
 Db 526 QKAKEQKAQLAOTLQOEQASQSGLRLHQVEQJQSSLQK---EQQLEEV-AEKQEATRQDH 581
 Qy 307 SLRERAYAAERQREQWEEKBOTLQFOQSKMACQLYREKYNALQAVQCELOKERDQAYS 366
 Db 582 A-QQLATAAE-BREASLRERDAALKOLE-----ALEKEKAALKI----LQQQLYANE 629
 Qy 367 ARDSAQEISQSLVEREDSLRROYFEL-----TDQCVCELTQRQLQAEPPG 412
 Db 630 ARDSAQTSVTQREKALSKVKEELQACVETARQOHEAQAAQVAEFLQRLSEQ---- 684
 Qy 413 VLKQEAATRECPREKORLYRMHAICPRDDSDCSLVSSTSOLLSISATSRELYDSFR 472
 Db 685 ---QKATEKERVQAQEKDQL-----QEQQLQALKESLKVTKGSLEBKR 723

Qy 473 SSSPAPPSSQSQSLYKRYAEDGEEPWSSSCLEIPEGDPGALPGAKAGDPHDYELLTDAD 532
 Db 724 RAADAEQOQRCISEURAE-----TRSLYEQHKERKEKEERACKGIEARLJOLGE 776
 Qy 533 LPQLESSLQPSPGRLLDVSSESVLMLRPARRILSQTMLAFQGDALLEQTSVIGNLGTG 592
 Db 777 AHQAETE-----VLRLELAAMAQHT-----AESCEQL----- 806
 Qy 593 IFTHRYTPGSAADOMALRPGTQIVMYDYEASEPLFKAVLEDITLLEAVGLLRRVDGFCLL 652
 Db 807 -----VKEVAAMRDGYE---DSQQEBAQYGMFQE-----LM 836
 Qy 653 SVKVNTDGYKRULQDLEAKVATSGDSFYIYRNLMABSGRAKGEIQLVHGENEVLHVTDTMFQG 712
 Db 837 TLKEECKSKAROELQAEKVA-GIEHSESELTSROONKLAEL--IRN----- 880
 Qy 713 CGCWHAHRVNNTYTMKDTAAHGTIPNYSSRAQQQLIALQDMTQCTVTRKPSSGGPQKLV- 771
 Db 881 ---LARALQQYQKEVYRAQKLADDLSTLQERMAATESKVARLETLYVRK--AGEQQTAS 934
 Qy 772 RIVSMIDAKASPLRLSFDRGQLDPSRME--GSSTCFWAESCLTLVPTYLWPHRPARP 828
 Db 935 RELVKEPARA-----GDROPENWLEEQGRQFC-----STQA 965
 Qy 829 PVLLYPRAVGKILSEKUCLLQGFFKKCIAEYLQEEYFAWSQRGDITQEGETSYGGRCWTR 888
 Db 966 ALQAMEREAEQNGNE---LERLRAALMESQQQERGQPREVALTQERG-----R 1015
 Qy 889 HAVESLMEKNTHALDVLQDVSCTLHMDIFI -VTHVSYNEKMAK-----KLK----- 936
 Db 1016 AQADLALEKAARAELAELVRNLQNALNEQREVFATQEAALAHALTEKEGRDQELAKLRGIEAA 1075
 Qy 937 --KGLQRGTSBELLROTQVKLKEOLAKKE 957
 Db 1076 QIKELEELROTQVKLKEOLAKKE 1098

RESULT 6
 US-09-152-294-1
 Sequence No. 6 Application US/09452294
 ; Patent No. 6287790
 ; GENERAL INFORMATION:
 ; APPLICANT: Le lievre, Sophie
 ; APPLICANT: Bassell, Mia
 ; TITLE OF INVENTION: UTILIZATION OF NUCLEAR STRUCTURAL PROTEINS FOR TARGETED
 ; TREATMENT AND DETECTION OF DISORDERS
 ; FILE REFERENCE: IB-1454 - Sequence Submittal
 ; Patent No. 6287790
 ; CURRENT APPLICATION NUMBER: US/09/452,294
 ; CURRENT FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: 60/110,420
 ; PRIOR FILING DATE: 1998-11-30
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 1
 ; LENGTH: 2101
 ; ORGANISM: Homo sapiens
 US-09-452-294-1

Query Match Score 5.18; Score 262; DB 4; Length 2101;
 Best Local Similarity 19.9%; Pred. No. 5.7e-13;
 Matches 208; Conservative 163; Mismatches 376; Indels 296; Gaps 39;

Qy 4 LCRDSLTADEETLWEMMESHRRIVRCICPSREITPYLRAKYVLCQLDDEEVLHSPLR 63
 Db 263 ILINEKQAASPLEPKELLELRDKN-----ESTLMRLHETLKQCDQDLTEKSQMDRK 312
 Qy 64 TNSMRAGHLIDLKLRKGNAIAF-LESFKHNPDVYTLYVGLQDPDVDFNSFSGLMETS 122
 Db 313 INQLE-----ENGDSFKLREFASHLQLQDALNELTEEHSAATQENWLEKQA 360
 Qy 123 KUTECLAGAISGLQEEENQGKEYLLRRCQOLQHGLAETRAEGLHOEADISMKR 182
 Db 361 QLEKELSAA---LQDKKCLE-KNEILOGKQSLEBHS-----QIQDNPPQEKG 406
 Qy 183 EYSAHPEVLRUKDEMISLS-----LHYSNALQKEKELAASRCRSLOQEELYL-- 228
 Db 407 EVLGDVYLQLETLIKQEATLAANNTQLQARVMLETGRQQAKLAEKGHEEEKQLOSS 466
 Qy 229 -----LKOELQRANMVSSCELLQEQSLRT-----ASDQEGSDEELNRK 268
 Db 467 LITDQSSISNLSQAKEELQASOAHGARLTAVASHTSETLNTNATQQ-DQEVLAGLK 525
 Qy 269 EE-----NEKLRSQSLDPLARGQELVERIH 306
 Db 526 QKAKEQKAQLAOTLQOEQASQSGLRLHQVEQJQSSLQK---EQQLEEV-AEKQEATRQDH 581
 Qy 307 SLRERAYAAERQREQWEEKBOTLQFOQSKMACQLYREKYNALQAVQCELOKERDQAYS 366
 Db 582 A-QQLATAAE-BREASLRERDAALKOLE-----ALEKEKAALKI----LQQQLYANE 629
 Qy 367 ARDSAQEISQSLVEREDSLRROYFEL-----TDQCVCELTQRQLQAEPPG 412
 Db 630 ARDSAQTSVTQREKALSKVKEELQACVETARQOHEAQAAQVAEFLQRLSEQ---- 684
 Qy 413 VLKQEAATRECPREKORLYRMHAICPRDDSDCSLVSSTSOLLSISATSRELYDSFR 472
 Db 685 ---QKATEKERVQAQEKDQL-----QEQQLQALKESLKVTKGSLEBKR 723

Query Match Score 5.18; Score 262; DB 4; Length 2101;
 Best Local Similarity 19.9%; Pred. No. 5.7e-13;
 Matches 208; Conservative 163; Mismatches 376; Indels 296; Gaps 39;

Qy 4 LCRDSLTADEETLWEMMESHRRIVRCICPSREITPYLRAKYVLCQLDDEEVLHSPLR 63
 Db 263 ILINEKQAASPLEPKELLELRDKN-----ESTLMRLHETLKQCDQDLTEKSQMDRK 312
 Qy 64 TNSMRAGHLIDLKLRKGNAIAF-LESFKHNPDVYTLYVGLQDPDVDFNSFSGLMETS 122
 Db 313 INQLE-----ENGDSFKLREFASHLQLQDALNELTEEHSAATQENWLEKQA 360

NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS: TESTA HURWITZ & THIBEAULT
; ADDRESSEE: TESTA HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06160
; FILING DATE: 19930621
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESO, EDMUND R
; REGISTRATION NUMBER: 7,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-06160-4

Query Match Score 5.18; Length 2101;
Best Local Similarity 19.9%; Pred. No. 5.7e-13;
Matches 208; Conservative 163; Mismatches 376; Gaps 39;

Qy 4 LCRDDSLTADEETLWEMMESHRRHTRVRCIPSRLTPYLRAKVLCQLDDEEVLSHSPRL 63
Db 263 LINEKQASPLEPKELERDKN-----ESLMRHLHETLKQDLKTEKSNDRK 312
Qy 64 TNSAMRAGHLLDKLTKGNGIAF-LESLKFHNPDVYTLVGLQPDVFNSNSGLMETS 122
Db 313 INOLSE -----ENGDSFKLREFASHLQLQDAINELTEHHSKATQEWLEQA 360
Qy 123 KTCI CLAGAIGSQQEENQKEVILLRRCQLOQHGLA TRA GHQ LEADHSRMKR 182
Db 361 QLEKELSA--LQDKKCLEE-KNEILQGKL-SQLEEHLS----ASDOESDEELNRLK 268
Qy 229 -----LKOELQRANMVSSCELEQEQSLR-----ASDOESDEELNRLK 268
Db 467 LTDLQSSISNLSQAKEELQASQAHARLTQAVASLTSELTTLNNTIQQQ-DQEAGLKL 525
Qy 307 SLRERAVA AERO EQYWEKEOTLLOFOKS KMACOL YREKYNALQAVCELOKERDAYS 366
Db 630 ANDSAQTSTVQDQREKAELSKRVEELQACVETAROHEAQAVAE-----QEQLOLQVANE 629
Qy 413 VLRQEARTREP PREKORLYRVMHAICPRDDSPCSLSVSSTESATSSRELVDTSFR 472
Db 685 --QKATEKERAQERDOL-----QEQLOALKSLKTVTGSLEEKR 723
Qy 473 SSSPAPPSQOSLSSQYKRAE DFGEPWSSFCLEPIPEGDPGALPGAKAGDPHDYELDTAD 532
Db 724 RADALEBQQRJLSEJKA E-----TRSLYEQHKRERKELBEERGKGLEARLLQLG 776
Qy 533 LDQLESSLQPVSPGRDVLVSSECVLMBRRPARPILSOUTMLAFQGDALLEQISVIGNLTG 592
Db 777 AHQAE TE-----VLRRELAEMMAQHT----AESECEQL-----806
Qy 593 IFTHRVTPGSAADQMALRPGTQIVMDYEASEPLPKAVLEDPTTLEAVANGLLRVDGFCC 652
Db 807 -----VREVAAMPDGE----DSQQEAEQYGMFQO-----IM 836
Qy 653 SYVKVNTPGYKRLLDQKAVATSGDFSYI RVNLAMEGRAKLQVHCNEVLHVTDTMFOG 712
Db 837 TKEECEKAROLQEAKEKVA-GIESHSELLSRQONKLAEL-HAN-----880
Qy 713 CGCWHAHRVNSYTMKDAAHGTIPNSRAQQQLIAJQDMHQCTVTRKPSGGPQKLV- 771
Db 881 ---LARALQQQKEVYRAQKADDLSTLQRMMAVTSKEYARLETIVRK--AGEQQETAS 934
Qy 772 RIVSMKDACKASPLRLSFDRGQLDPSRM--GSSTCEWAESOLTLYPTVTLWPHRPARP 828
Db 935 RELVKEPARA-----GROOPENEEQGROFC-----STQA 965
Qy 829 PVLLPVRAVGKTLSKEKCLLQFQKCLAYLSQEEYEAWSGRMGTQI QEGEVSGGRCWVTR 888
Db 966 ALQAMEREAEQMGNE---LFLRRAALMESQQQBERGQOERREVARTQERG-----R 1015
Qy 889 HAVESLMEKNTHALDYOLDSYCTLHRMDIPI--VTHVSYNEKMAK----KLK-----936
Db 1016 AGDALEKAARAELENRLQNLNEQRFETLQEA LAHTEKEGKDQELAKL RGLEAA 1075
Qy 937 -KGLOQIGTSESEQLIEAARQEE 957
Db 1076 QKAELEELRQTVKQLEQKARKE 1098

RESULT 7
; Sequence 4, Application PC/TUS9306160
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
US-08-353-700-1

Query Match 4 6%; Score 338; DB 1; Length 3248;
Best Local Similarity 19.4%; Pred. No. 1.3e-10;
Matches 214; Conservative 192; Mismatches 393; Indels 306; Gaps 48;

Query 2 GELCRRDSALTALD-----EETLWEMMESHRIURVCICPSRLTPYLRAKYVLCOLDE 54
Matches 214; Conservative 192; Mismatches 393; Indels 306; Gaps 48;

Query 2061 GELDTMSKTKTFLQSEKMEKFT - QELLESHQSCLHC1-----QVAE 2102
Db 837 TLKEECARQELQAEKVA - GIESESELQISRQNKLAEI - HAN----- 880

Query 55 BEVLHSPLRTNSAMRAGHLDDLLKTRGKNGAIAFESLKFPHNPDVYTLVGLQPDVDFSN 114
Db 2103 AEVKETKELIQLT - SSDYSELLKD-----KTH - LQEQLSLEKD----- 2139

Query 115 FSGMMETSKITTECLAGIALGSQEEENQGKEVJLRRCQLQELG----- 161
Db 2140 - - - SOALSITKC - - - ELENQIAQLKEKELVKESESLSQARLSESDYEKLNVSKAL 2189

Query 162 - - - - - LAETRAEGLHOLEADHSRMKREYSA - - - - - HFHEYLR-----L 194
Db 2190 EAALVEKGEPALARSTSQEE - VHQLRGTEKLRVRIEADSKKQHTAEKLUKERENDSL 2248

Query 195 KDEMILSLHYSNALQEQKELAASCRSLQEELYLKKQELQRANVSSCEL - ELEQSLRT 253
Db 2249 KDKVENLERQLOMSEENQELVILDAENSKRAEVETLKTOB - ENARSLLKFIELDLVTLRS 2306

Query 254 ASDO - - ESGDEELLKKEEIKLSSLTSIAEKDL - OSLDEAEGSRQFVERHSCRE 310
Db 2307 EKENITKQIOEKOGOLSELDKLSSKFSLSLEKEAQEIQKEESTAVEMQNQKELNE 2366

Query 311 RAVAERQREQQYEEKWEQT-----LQOFQSKMACQOLYREK----- 346
Db 2367 - AVALCGQEIIMKATEQSLDPPEEEQRLRNLSEKLRDAEDEKKQQLVQLIKESEH 2425

Query 347 - VNALQAQCYCELQKRDQAYSARDSAORETSOSLVEKDSLROYFEELTDQVCELRQLRQ 405
Db 2426 HADLJKGRVNLERELEIAINTQEHAALEAENSKEGEVETLKAKEGMTQ - - - - - SLRG 2478

Query 406 LQAEPPGVKQEARTRREPCKRQLVRNHAICPRDDSDCSLVSTTESQQLSDISATSSR 465
Db 2479 LEVDVTIRSEKENLTNEQKEQRI - - - - - SELEIINSSENFILEQ - - - - - KEQ 2523

Query 466 ELVDSFRSSSPAPPSSQSLYKRVAAEDFGEEPWFSSCLEIPGDPGALKAGDPHLDY 525
Db 2524 EKVOMKEKSSTAMENLQTLKELNVRVAALHNDQAC - - - - - KAREQNLSS 2569

Query 526 EL - - - LDFTADLPQ - - - - - LESS - - - - - LQPVSPQRDVSSEGVLMRRPARRIL 566
Db 2570 QVECIELEAQLLQGLSDEAKNNYIVLQSSSKGLQEOVEDGQKLEK - - - - - KDEEISRLK 2624

Query 567 SQVIMLAFOGDALDQISITG - - - - - NLTGJFIRHPTPGSAADQMLRPGT 613
Db 2625 NQIQ - - - DQEQLQSKLSQVEGEHOLWKONELRNLT - VLEOKIQIVLOSKNASLQDTL 2679

Query 614 QIVMVDEASEPLPKAVLDTITLEAVGGLJRRDGFCCLSVVKNTDGYKRLQDLEAKYA 673
Db 2680 EVLQSYKNE - - - NELET,KMDK - MSFEVKVMTAKETELOREMHEMAQTAELQEE 2734

Query 674 TSGDSFYIRVNNAMEGRAKGELQVHCNEYLVHTDMFQ - - - - - GGCWHAHRVNS 723
Db 2735 LSGE - - - - - KNRLAGELQILLEEIKSKQDQKLTELENSELKKSLDCKMKDQVE - 2783

Query 724 YTMDTAAHGTIPNS - - - RAQOQILALIQDMHQCTV - - - - - TRPKSSGGPQKL 770

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 COMPUTER TYPE: PC-DOS/MS-DOS
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/353,700
 FILING DATE: 09-DEC-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: REED, JANE T.
 REGISTRATION NUMBER: 36,252
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 563-4100
 TELEFAX: (215) 563-4044
 INFORMATION FOR SEQ ID NO: 1:
 LENGTH: 3248 amino acids
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3248 amino acids

Qy	55	EYVLHSPLRTNSAMRAGHLLDILKTRGNGAIAFLESLKFKHNPDVYTLVGLQDDDFSN	114
Db	2103	AEVKETTELLQTL - SSDVSSELLKD -	KTH - LQEKLQSLKD -
Qy	115	FSGLMETSKEULGAGTSLOEFLNQEKGOKEVLLRRCQOLOEHLG -	----- 161
Db	2140	---SQAUSLTKC- ----- ELEQIAQNKEELVKESESLSQARLSEEDYEKUNVSKAL	2189
Qy	162	-----LAETPRAEGLHOLEADHSRMKREVA -----HEHEVLR-----L	194
Db	2190	EAALVERGEFAEFLRSSTOOE - VHQLRGTRGTEKLVRVIADEKKOLHIAEKLERENDSL	2248
Qy	195	KDEMILSLSLHYNSALQKEELAASCRSLOEYLKLQELQRANNSVSCCL - ELOEQSLRT	253
Db	2249	KDKVNJERQMLSEENQELVILDAENSKAEVETLTQTL - EMARSLKIFELDLVTLRS	2306
Qy	254	ASDQ - ESGDEAELNRKBENEKLRSLSLAEDKILE - QSLDEAEGRSROELVERIHSRLRE	310
Db	2307	EKENLTHIQEKGQSLSEDKLSSPKSLSLEEKEQAOIKEESKTAVEMLNQKELNE	2366
Qy	311	RAAAEROREQYEEKQTLL -----LQFQKSXMAQOLYREK -----3446	
Db	2367	-AVAALCGDOETIMKATEOSLDPPTEEHOLRNSTEKLARLEADEKKQLCVQLQKSEH	2425
Qy	347	VNALQAQCVOCELOKERDAYSARDSAQREREISQSLVKEKDSLRQVFELTQVCELRQLRQ	405
Db	2426	HADLKLGRVNLERELLEARTNQEHAALEAENSKGEVEETLKALEGMTQ -----SLRG	2478
Qy	406	LOAEPPGVLKQEARTREPORCPREKORVMAICPRDSDCLSYSTESQDLSLDSATSSR	465
Db	2479	LEIDVVITSEKENNLNTLQKEQERI -----SELEIINSSSENILQE -----KEQ	2523
Qy	466	ELVDSFRSSSPAPPSQOSLYKRYADEFGEPEWPSFSSCLEPIEGDPGALPGAKAGDPPHLDY	525
Db	2524	EKVQMKEKRSSTAMEMLQTKELNERVAALHNDQEAC -----KAKEQNLS	2569
Qy	526	EL --- LTDADLPQ --- LESS --- LQPVSPGRDVSSEGYLMRRRPARRIL	566
Db	2570	QVECLEKAQQLGLDEAKNNNTVLAQSSVKGLIQEVEDGOKGLEK -----KDEEISRLK	2624
Qy	567	SOVTMIAFGDALLEQISIVIGG -----NLTGIFIHRYTPGSAADQMALRPGT	613
Db	2625	NQIQ --- DQEQLVSKLSQVEGHQWLKEQNLEURNLT - VELEKQIVQLOSKNLSQDTL	2679
Qy	614	QIVMVDXEASEPLFKAVLEDTTLFEAVGULLRRTYDGFCCLSVVKVNTDGYKRLLQDLEAKVA	673
Db	2680	EVLQSSYKNLE --- NELELTMDK - M\$FVEVKNMKMTAKETOLEREMHENAQKTAELQEE	2734
Qy	674	TSGDSFYIRVNLAEGRAKGELQYHCNEVLYHVDTMQ -----GGCGCWIAHRVNS	723
Db	2735	LSGE ----- KNLRLAGOLLIEELKSSKDQOLKEELTLENSELKKSLDCMHQDQE	2783
Qy	724	YTMKDTAAHTPINYS --- RAQQQLALIQDMTQCTV -----TRKPSSSGGPQKL	770
Db	2784	---KEGVYREIAEYQQLNHEAEKKHQALLDTNQKEYEVIQTYREKLTSKKEECLSQKL	2840
Qy	771	VRIVSMKAKASPLRLSFD ----- RGOLDPSRMEGSSSTCFWAESCLTVPYTLWPHR	823
Db	2841	---EIDLKSSKEELNNSLKAATTQILEELKKTMD -----NIKYVNQ	2879
Qy	824	PAPRPVILIPRAYERK1 - LSEKLC - LLQGFFKKCLAEYLQEEYAWQSQRDIIQEGEVSG	881
Db	2880	LRKEN - ---ERQGMKUILLIKSCKQLEPEKEILQKELSOLQAQEQKQTGTVMTDKVDE	2934
Qy	882	GRCWVTRIAVESLMEKHNHALDVQLDSVCTLHMRDIFLIVHVSNEKMAKKLKGRLR	941
Db	2935	LTTEI - KELKETLEEKTEA - DEYLKXCSL -----LISH ----- EKLEKAKEM	2976
Qy	942	LGTSEEQLEAARQEFGDLRACP1	966
Db	2977	LETQVAHL - CSQSKQDSRGSPLL	2999

```

RESULT 10
US-08-328-254-6
Sequence 6, Application US/08328254
Patent No. 5710022
GENERAL INFORMATION:
APPLICANT: Zhu, Xueliang
APPLICANT: Lee, Wen-Hwa
TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4310 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328, 254
FILING DATE: 24-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/141, 239
FILING DATE: 22-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31, 815
REFERENCE/DOCKET NUMBER: P-CJ 1191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-9949
TELEFAX ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2482 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-328-254-6

Query Match 4.6%; Score 237; DB 1; Length 2482;
Best Local Similarity 19.4%; Pred. No 1e-10;
Matches 214; Conservative 192; Mismatches 393; Indels 306; Gaps 54
Qy 2 GEICRDRSALTA[D-----ETLWENMEHSHRRIVACICPSSLTPYRQAVALCQLDE 54
Db 1333 GEDLDMSKTTA[DQLSERKMF-----QELBOSH[CLCI-----QVAE 13
Qy 55 EEVLHSPLTNSRAMGHLIDLKTRKGNGIAIFLESLIKHNPDVYTLYTGLOPDVDFSN 11
Db 1375 AEVKETTELLQT[L--SSDVSEL[KD-----KTH---LQEQLQSLEKD----14
Qy 115 FSGGMETSKRLTECLGAGLSQEELNQENGQKVEYLRLRQQLOBHGLG-----16
Db 1412 ---SQALSLTKC----ELENQIAQNKEKKVKESESLSYDYEKLNVSKAL 14
Qy 162 ---"LAETRAEGLHQLEADHSRMREVSA----HFHEVLR----L 19
Db 1462 EAALVEKGKFALRSSTQEQE-VHQLRGTEKLVRIEADEKKHIAEKLKERERENDS 15
Qy 195 KDEMILSLSHNSNQERKLAASCRSLSOEE[LKQELQRANVVSSCBL-EHQEQSLRT 25
Db 1521 KDKVENLRLQSEENQELVILDAENSKAEVTLKTQIE--EMARSUKVFEELVTLRS 19
Qy 254 ASDO--ESGDEELNRKKEENEKLRSLTSFLAEKDILE-OSLDPARGSRQELVERBHSRLE 31
Db 1579 EKENLTQIQQEQLQSLDLSKLFKLSSEKEQAEIQKEKSATAVMQLNQKLNE 16
Qy 311 RAYAERQEQYEEKEQT[L--34
Qy 311 LQFQSKNACOLYREK----LQFQSKNACOLYREK----34

```

RESULT 11
US-08-585-576-1
; Sequence 1, Application US/08685576
; Patent No. 5906819
; GENERAL INFORMATION:
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Nakano, Takeshi
; APPLICANT: Ito, Masaki
; APPLICANT: Takahashi, No. 5906819uaki
TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/685,576
 FILING DATE: 24-JUL-1996
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-325129
 FILING DATE: 20-NOV-1995
 PRIOR APPLICATION NUMBER: JP 8-17150
 FILING DATE: 05-JAN-1996
 PRIOR APPLICATION NUMBER: JP 8-131206
 FILING DATE: 26-APR-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 TELEPHONE: (202)672-5300
 REFERENCE/DOCKET NUMBER: 29, 768
 TELEX: 904136
 TELECOMMUNICATION INFORMATION:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1388 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-685-576-1

Query Match 4.2% Score 216.5; DB 2; Length 1388;
 Best Local Similarity 19.6%; Pred. No. 2.3e-09;
 Matches 151; Conservative 129; Mismatches 292; Indels 199; Gaps 28;

Qy 6 RDSALTADEE-TLWEMMESHRRHIVR--CICPSRLTPYLROAKVLQLDBEEVLHSPR 62
 Db 476 RLEVAKELDEETLTKNEVESTLRQLEREKKLHHKNAEYQRKA-----DHEADKRN 528

Qy 63 LTNSAMRAGHLDDLIKTRKGNGATAF-----88
 Db 529 LENDVNSLKDQLEDLKRNQNSQISTERKVNQLRQLDETNNLLRTESDTAARIKRTQAE 588

Qy 89 --LESLIKHPNDVY---TLWGT-LQPDVDFSNSGMLET-----SKLTTECLAGAI 32
 Db 589 SKIQOLQLESNNRDIQDKNCNLLEPAKLKLEKEFLNLSQVLESERDRTHGSEILNDLOGRI 648

Qy 133 GSLOPDELNOEGKQKEVLLRQC----OQBHQHGLAETRAEGL----HQLEADHSHRMRE 183
 Db 649 SGLEDVKNK---ILLKVBLERKQLOQERTDLEKERNMNEIDMTQTLKVHQQSLEQ 704

Qy 184 VSAHFHEVLRKDE--MLSFLSHNYSALQKELAASCRSRLQOEELYLKQELQRANMVS 240
 Db 705 ETEHKATKARADKNIYBSEIAKSEAMKEMERKLSEERTLQKVENNILLEAKRCISL 764

Qy 241 SCELLEQESLR-----TASDQESDEELN-----265
 Db 765 DCIDLJKQSOKINELLKQDKVLYNEDVRNLTFLKIEQETOKRCLTQNDLKMQTOQVNTLKMSE 824

Qy 266 -RLKBEENKURSLTSFLSAEKD-I-LEQSLDEARSQRLQELVERTHSRLRERAAERQEQYW 323
 Db 825 KQLQRENHNHLEMMSLEQNAEFLQKERAQKRDGMKELQDQ-----EAEQYF 872

Qy 324 EKEEQTLLOFOQSKMACQLYREKYNALQAQOCVCELOKERDAGMKEQDQ-----
 Db 873 STLYKT--QVRELKECEEKTKLCKELQKQELQDERSL----AAQLEITLTKADE 925

Qy 384 SLRROYFELTDQCELTQRLQQAEPGVYKOFARTREPCCPREKORLYRMHAICPRDD 443
 Db 926 QLRSIAE-----EQYSDLEKE-----IMKELEIKMARRHKQELT-----ERDA 966

Qy 444 DCSLYSSTESOLSDLIS--ATSSRELVDSSFRSSSPAPPSSQOSLYKRVAAEDFGEPWSFSS 501
 Db 967 TIASLBETNRUFLTSVANLANEKEBNNUKIRE-----AEBQLSPLKDBEI-----1011

Qy 502 CLEIPPCGDPGALPGAKAGDPHDYELL-----TADLPQLESSLQPVSPG-----R 547
 Db 1012 -----SAAAIIKAQFEKQLLTERTLTKTQAVNKLAEIMRKVEVKRGNDTDFRR 1058

Qy 548 LDVSESGVLMRRPARRILSXYMLAFQGALLEQISVIGGNLTGIFHVRVTPGSAADM 607
 Db 1059 KEKENRKHLMEKLSEREKLQ-QMICKYQELNEMQAQAESESQIRIELQMTLDSKSDIE 1117

Qy 608 ALRPCTQIOMYDYEASEPLPKAVLEDTTLEAVGLLRRVDFCCLSVKNNT 658
 Db 1118 QLRSQLOALHIGLDSS--LGSGPDTADGFPE-SRLEGWLSLEPVNNT 1165

RESULT 12
 US-09-104-324B-4
 ; Sequence 4, Application US/09104324B
 ; Patent No. 623460

; GENERAL INFORMATION:
 ; APPLICANT: Treci, Ozlem; Sahin, Ugur; Pfreundschuh, Michael
 ; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,
 ; And Methods For Identifying Pathogenic Markers In A Sample
 ; TITLE OF INVENTION: No. 6232460mial Cells
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fulbright & Jaworski LLP
 ; STREET: 666 Fifth Avenue
 ; CITY: New York City
 ; STATE: New York
 ; ZIP: 10103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
 ; COMPUTER: IBM
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/104, 324B
 ; FILING DATE: 25-June-1998
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/892,702
 ; FILING DATE: 15-July-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 623460man D.
 ; REGISTRATION NUMBER: 30, 946
 ; FEEPAID NUMBER: LUD 5491
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 318-3000
 ; TELEFAX: (212) 752-5958
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 976 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-09-104-324B-4

Query Match 4.1% Score 210.5; DB 4; Length 976;
 Best Local Similarity 22.8%; Pred. No. 4.2e-09;
 Matches 119; Conservative 76; Mismatches 189; Indels 139; Gaps 20;

Qy 12 TALDEETLWEMMESHRRHIVRVCIPSRLTPYLRAQK-----L 49
 Db 303 TKLQSENQKQSTEKQHH-----ITKLEEDIKVKLQSRSTVQALEEDQIATKTI 352

Qy 50 COLDEEE-----VLHSPLRTNSAMRAGHLLDKTRGKNGAIAFLESLKFINPD 98
 Db 353 CQLTEEEKTQEMESENKAARAHSFVVTETTCVSLLEELRTE-----QQRLEKEDQ 404

Qy 99 VYLVTVGQPDVDFSNFSGLMETSKLTCLAGATGSLOELNQEKGOKEVLLRRCQQL-- 156

Best Local Similarity 20.8%; Pred. No. 9.3e-09;
 Matches 105; Conservative 105; Mismatches 184; Indels 110; Gaps 20;
 QDy 40 TPLRQAKVLCQLDDEEVYHSP--RLTNSAMRGHLDDLKTR--GKNGAIAFLE-- 90
 Db 361 TKALSLALAR----EEFLAKEELERTNMKLA--EMEDVSSKRDVGKN--VHLELSK 412
 QDy 91 -SLKFHNPDVTYLTVGLQDPDFSNEGLMETSKLTECLAGAIGSLOEELNQEKGKEVIL 149
 Db 413 RALETQMBEMKTOQLEELDQLASEDAKL---RLEVNNQALKGFERDQARDQNE- 466
 QDy 150 LRRQCQLOEHIGLAETRAEGLHQLEADHSRMKREVAHEVLBLKDDEMUSLHYSNAL 209
 Db 467 -EKRLQRLQRLHEYET-----ELEDERNERALAAAARK--KLEGDLRDLQADSAI 515
 QDy 210 QEKEALAASRCSRSLQEEYLKLQELQ-----RANMVSSCELEQEQSL 251
 Db 516 KREEEAIKQLRQLQAMKDFQRELEDAARSDREFATAKENERKAKSLEIDLQLOQEDIA 575
 QDy 252 RTASDQESGDEELNRKEE-----NEKLRSLTSFLAEAKDLEQSLDEARGSRQE 300
 Db 576 AAEARKOADELERKELABEABLASSLSGRNALQDERRRLARIAQ--LEEFLLEEQQNMEA 632
 QDy 301 IVERIHSLRERA----VAAEFQRQEWKEKEATLQFQSKRMACQI-----YR 344
 Db 633 MSDVRVKATOQAEQNSLNELATERSTAQKNESSARQO-LERONKELRSKLIHEMEGAVSKFKF 691
 QDy 345 EKVNALAQCELCIQLERDOAYAARDSAQOREIQSISLVEKOSLRRQVFELTDQVCELTQLR 404
 Db 692 STIAALEKKIAQLEQVQEARPKAA---TKSLQKOKRKLLKEULLQEDERKAEOQY 747
 QDy 405 QLQAEGPPGVLKQFARTREPCKPREKRQLYRVMHAICPRDDSPCSLYSSTSOSLSDL-SATS 463
 Db 748 EQAEKGNAVRKQKLRQLEBAAEESQR-----INANRKLQRELDEATE 790
 QDy 464 SREL---VDSFRSSSPAPPSSQSQS 483
 Db 791 SNEAMGREVNALKSKLRGPPPQET 814

RESULT 14
 US-08-742-923A-6
 Sequence 6, Application US/08742923A
 Patent No. 5869611
 GENERAL INFORMATION:
 APPLICANT: Liu, Pu
 APPLICANT: Collins, Francis S.
 APPLICANT: Siciliano, Michael J.
 APPLICANT: Claxton, David
 TITLE OF INVENTION: Markers for Detection of Chromosome 16
 TITLE OF INVENTION: Rearrangements
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
 STREET: P.O. Box 828
 CITY: Bloomfield Hills
 STATE: MI
 COUNTRY: USA
 ZIP: 48303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08742-923A
 FILING DATE: No. 5869611ember 1, 1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, DeAnn F.
 REGISTRATION NUMBER: 36683
 REFERENCE/DOCUMENT NUMBER: 2115-00869DVC
 TELECOMMUNICATION INFORMATION:
 NAME: Smith, DeAnn F.

TELEPHONE: (810) 641-1600
 TELEFAX: (810) 641-0270
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 816 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 S-08-742-923A-6

Query Match	Best Local Similarity	Score	DB 2;	Length	816;
	Matches 105; Conservative 105;	Pred. No.	9.3e-09;		
	Mismatches 184;	Indels	110;	Gaps	20;
Y	40 TPYLQRQAVKVLQCOLDEEEVYLHSP--RLTNSAMRAGHLLDLIKTR--GKNGAIAFLE-- 90				
D	361 TRALSLARL--EPAAKEELEPTNRMKKA-EMEDLYSSKDDVGKN--VHEDELSK 412				
Y	91 -SLKFHNPDVYTIVLTGLOPDYDFSNFSGLMETSKITECLAGAIGSLOEELNQEKGKGEVL 149				
D	413 RALETQMEBKKVQLEELDEDQASEDAKL--RLEVNMQAUKQPFERDQLQARDERNE-- 466				
Y	150 LRRCQOQEHGLAETRAEGLHQLEADHSRMKREVSASHFHLVRLKDEMMSLISLHVSNAL 209				
D	467 -EKRRQLORQHEYET-----ELEDERNERALAAAKK--KLEGSDLKLELQADSAI 515				
Y	210 QEKELAASRCSRSLQEEILYLKQBLQ-----RANMVSSCBELEQDQLS 251				
D	516 KOREATAKQLRQLQAMQKDQRELEDAKSREIFATAKENKKKANSLEADMQGEDIA 575				
Y	252 RUDSQESEDEELNRKEE-----NEKLRSLTFSLEAKDILEQSLDEARGSROE 300				
D	576 AAERKQADLKEELAEELASSLSSGRNLQDOKRLEARIAQ---LEEELEEEQGNMEA 632				
Y	301 LVERIHSRLERA-----VAARQREQWEEKBETLQFQPKSMATOL--YR 344				
D	633 MSDRVRKATQQAEQLSNELATERSTAGKNEARSQQ-LERQNKELSKLHMEGAVSKFK 691				
Y	345 EKVNALQQVCLQKEDRQDAYARDSAQREIQSOLSYVEKDSLRQVFELTDQVCELTQLR 404				
D	692 STIANLERKIALEQVPEQAREKQAA--TKSLRQDKKKKEILQVETERKMAEQYK 747				
Y	405 QLQAEPCCPVLAQKQEARTEPCPREKORLYRMHAICPRDDCSLVSSTESQQLSDL-SATC 463				
D	748 EQAEKGNAWRKQVKQLEBEEBSQR-----INANRKLQRELDDEAT 790				
Y	464 SREL---VDSFRSSRAPPSQ 483				
D	791 SNEAMGRENVALNSKLQRGPPQET 814				

RESULT 15
S-08-533-306A-4

Sequence 4, Application US/08533306A
 Patent No. 5837457

GENERAL INFORMATION:

APPLICANT: Liu, Pu

APPLICANT: Collins, Francis S.

APPLICANT: Siciliano, Michael J.

APPLICANT: Claxton, David

TITLE OF INVENTION: Markers for Detection of Chromosome 16 Rearrangements

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 838

CITY: Bloomfield Hills

STATE: MI

COUNTRY: USA

ZIP: 48303

COMPUTER READABLE FORM:

DATA

SEQUENCE

MEDIUM TYPE: Floppy disk